HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Src-Homology 3 (SH3) Domains in Cellular and Virus-Host Protein Interactions

Constanze Schmotz

Department of Virology
Medicum
Integrative Life Science Doctoral Program
University of Helsinki
Finland

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Medicine of the University of Helsinki, Lecture Hall 2, Haartmaninkatu 3, Helsinki, on the 20th of December 2019, at 12 noon.

Helsinki 2019

Supervisor: Kalle Saksela, MD, PhD, Professor
Department of Virology
Medicum
University of Helsinki

Reviewers: Varpu Marjomäki, PhD, Docent
Department of Biological and
Environmental Science, Nanoscience Center
University of Jyväskylä

Marko Pesu, MD, PhD, Professor
Faculty of Medicine and Health Technology
University of Tampere

Opponent: Petri Susi, PhD, Docent, Adjunct Professor
Institute of Biomedicine
University of Turku

ISBN 978-951-51-5718-8 (paperback)
ISBN 978-951-51-5719-5 (PDF)
Picaset Oy
Helsinki 2019

To my family

Table of Contents

List of original publications

I. Kazlauskas A*, **Schmotz C***, Kesti T, Hepojoki J, Kleino I, Kaneko T, Li SS, Saksela K.: “Large-Scale Screening of preferred Interactions of Human Src-Homology-3 (SH3) Domains using Native Target Proteins as Affinity Ligands”, Mol Cell Proteomics. 2016 Oct;15(10):3270-3281.

II. Ylösmäki L, **Schmotz C**, Ylösmäki E, Saksela K.: “Reorganization of the host cell Crk(L)-PI3 kinase signaling complex by the influenza A virus NS1 protein”, Virology. 2015 Oct;484:146-52

III. **Schmotz C**, Uğurlu H, Vilen S, Shrestha S, Fagerlund R, Saksela K.:”MC159 of molluscum contagiosum virus suppresses autophagy by recruiting cellular SH3BP4 via an SH3 domain-mediated interaction”, J Virol. 2019 Mar 6. pii: JVI.01613-18. doi: 10.1128/JVI.01613-18.

*equal contribution of the authors

Abstract

Around 30 years ago, the discovery of modular protein domains began to change the scientific world's view on cell signalling. The multitude of ground-breaking discoveries in this field over the past three decades increased fundamentally the understanding of cellular signal transduction and how proteins transmit extra- and intracellular signals in this context. The SH3 domain was one of the first described modular protein domains and was soon found to be involved in numerous important cellular functions and signalling cascades. Moreover, SH3 domain-mediated interactions were also found to be hijacked by pathogens, especially viruses, which developed highly specific proteins mimicking cellular SH3 domain ligands. This molecular mimicry allows the virus to interfere, manipulate and exploit host cellular signalling for its own benefits and survival.

The vital elements of this doctoral thesis are the identification and characterization of so far unknown, high-affinity SH3 domain-mediated interactions and in the context of pathogenic SH3 domain ligand origin, also the elucidation of functional consequences for the host cell. The first part describes a large-scale screening approach uncovering high-affinity interactions between human SH3 domains and potential native target proteins. The identified robust SH3-mediated interactions are novel and thereby provide a valuable resource for future research. The second part extends the focus to SH3-mediated interactions between viruses and their host cells. These interactions are of exceptional impact since they enable the virus to manipulate or hijack the host's immune defence system and thereby promote viral survival and propagation. Two examples of viral proteins binding via an SH3-mediated interaction to host cell factors are presented in this thesis: - the interaction of influenza A virus (IAV) NS1 protein with host cell Crk-adaptor proteins in context of reorganizing the Crk(L)-PI3 kinase signalling complex and – the interaction between molluscum contagiosum virus (MCV) MC159 protein and host cell SH3 binding protein 4 (SH3BP4). The study addressing the reorganization of the Crk(L)-PI3 kinase signalling complex by NS1, focuses on structural and biochemical aspects of a newly discovered trimeric protein complex and the related functional changes in downstream signalling, such as Akt activation. The work presented on MC159 identifies host cell SH3BP4 as a new target protein of MCV and shows that the SH3-mediated interaction between MC159 and SH3BP4 is essential for suppression of cellular autophagy by MC159.

In summary, the work presented in this thesis highlights the fundamental importance of SH3 domains in the cell and its crucial role in the exploration of the host cell by viruses. By understanding cellular protein networks involving

SH3-mediated interactions and viral strategies targeting the same, new potential therapeutic targets can be identified, leading the way to novel or improved treatments.

Abbreviations

AAT	amino acid transporters
Abl	Abelson
AIDS	acquired immunodeficiency syndrome
AP-2	activating protein2
BCR	Breakpoint cluster region
CMA	chaperon-mediated autophagy
CPSF30	cleavage and polyadenylation specificity factor 30
cRNA	complementary RNA
Crk	CT10 regulator of kinase
CrkL	Crk-like
C-terminal	carboxyterminal
DAP1	death-associated protein 1
DD	death domain
DED	death effector domain
DISC	death-inducing signalling complex
DNA	deoxyribonucleic acid
ED	effector domain
EH	Eps15-Homology
$EspF_U$	F-like protein encoded on prophage U
FGFR	fibroblast growth factor receptor
FLIP	FLICE (FADD-like interleukin-1 beta-converting enzyme, aka caspase-8)-inhibitory proteins
HA	hemagglutinin
HIV-1	human immunodeficiency virus 1
HSV-1	herpes simplex virus-1
GIPC1	GAIP C-terminus-interacting protein
GRB2	Growth factor receptor-bound 2 protein

IAV	influenza A virus
IFN	interferon
IKK	IκB kinase
IRF3	interferon response factor 3
IRTKS	insulin receptor tyrosine kinase substrate
ISG	IFN-stimulated antiviral genes
JNK	c-Jun N-terminal kinase
KSHV	Kaposi's sarcoma-associated herpes virus
LC3	light-chain 3, aka Atg-8
M1	matrix 1
M2	matrix 2
MCV	molluscum contagiosum virus
mTORC1	mammalian target of Rapamycin complex 1
MYO6	myosin 6
NA	neuraminidase
NEMO	IKK subunit IKKγ
NEP	nuclear export protein
NF-κB	nuclear factor 'kappa-light-chain-enhancer' of activated B-cells
N-P-F	tripeptide-which is considered to be the core binding site of EH domains
NS	non-structural protein
OAS	2'-5'-oligoadenylate synthetase
OPT	optineurin
p85	regulatory PI3K subunit
p110	catalytic PI3K subunit
PA, PA-X	IAV polymerase subunit
PB1, PB2	IAV polymerase subunit

PI3K	phosphatidyl-inositol-3 kinase
PKR	dsRNA-dependent serine/threonine protein kinase
PRR	pattern recognition receptors
PTK	phosphotyrosine kinases
RBD	RNA-binding domain
RIG-I	retinoic acid inducible gene I
RNA	ribonucleic acid
RNP	ribonucleoprotein
RTK	receptor tyrosine kinase
SH2	Src Homology-2
SH3	Src Homology-3
SH3BP4	SH3 binding protein 4
Sos	son of sevenless
TBK1	TANK-binding kinase 1
TfR	transferrin receptor
TNF	tumour necrosis factor
TRIM25	tripartite motif-containing protein 25 (TRIM25)
VAR	variola virus
VV	vaccinia virus
vRNA	viral RNA
Y2H	yeast-two-hybrid

1. Review of the literature

1.1. Introduction to protein interactions

Proteins are masterpieces of evolution. From the simplest forms of life to the most complex organisms, nearly all biological processes depend on proteins and their execution of biological information. Depending on their amino acid sequence, structural features and function, proteins can be organized into different families and superfamilies, in which each member has related, but not identical functions. The architecture of most mammalian proteins resembles a “building block system”, in which different, independently folded and organized parts have distinct biochemical functions. These so-called "domains", either carry out certain biochemical reactions or mediate physical proximity between biological molecules. The principle of combining individual domains with distinct functions in one protein allows high diversity in the landscape of protein functions, as well as a high degree of regulation and control. Another advantage of this combinatory concept is the thereby limited size of needed genetic information, which is a favourable criterion for evolution [1].

Also, protein domains have a fundamental function in the process of global protein folding. By folding independently, they minimize possible interactions between amino acids within the protein. Furthermore, the initial folding of individual domains buries aggregation-prone hydrophobic amino acid residues in the core, while exposing hydrophilic residues [2, 3].

Complex cellular reactions and signalling cascades are carried out by multiprotein complexes and protein networks, respectively. Regarding the former, the presence of individual protein domains promotes higher degrees of stability for multiprotein complexes while reducing the likelihood of opportunistic interactions. For the latter, the combination of protein domains in the signalling cascade members determines the cascade’s assembly in terms of localization, direction as wells as control [1].

Detailed information on the protein structure facilitates understanding the biological function and importance of a protein. However, obtaining structural information by, e.g. nuclear magnetic resonance (NMR) or X-ray crystallography from full-length proteins is often challenging. In contrast, obtaining the information from isolated domains is mostly successful, allowing conclusions on the overall protein structure by solving the structure of individual protein domains.

For understanding the role of a protein plays in a cellular process, it is essential to identify and characterize its interacting partners and cellular environment. For this purpose, yeast-two-hybrid screens (Y2H), protein-arrays, and phage display screenings are most commonly used [1].

1.2. Phage display

"Phage display is one of the oldest and most powerful combinatorial biology methods" [4]. By fusing a polypeptide to the coat proteins of a bacteriophage (e.g. p3 or p8 of M13), it is possible to display the former in its native form. The DNA-encoded genetic information of the displayed polypeptide is located on a phagemid, which allows the generation of enormously large polypeptide libraries (>10^{11}) using basic molecular biological methods. Phages are optimal tools for displaying this type of libraries, since they are produced quickly in very high numbers and can easily be concentrated into small volumes. The protein of interest is usually immobilized when presented to the phage library (so-called "panning"), what allows easy access to the enriched phage population/s, which is/ are present if an interaction takes place between the protein of interest and the displayed polypeptide. Since every bacteriophage contains the DNA-sequence of the displayed polypeptide, it is possible to identify enriched clones by standard Sanger-sequencing [4]. Polypeptides displayed by the bacteriophages can be ligands for protein modules or the protein modules themselves when using the reverse approach. For identification of protein domain interacting partners applies the same principle of either displaying the domain ligand or the domain itself. The SH3 and WW (tryptophan-tryptophan) domain, but also single-domain antibody libraries are examples of domains that have been successfully expressed in and displayed on bacteriophages [1].

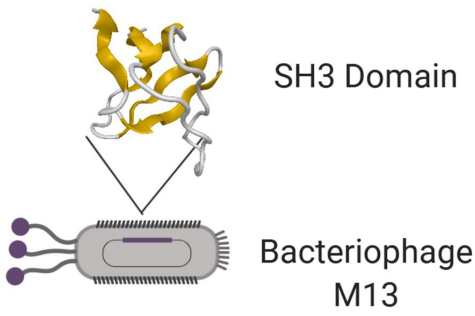

Figure 1. SH3 domain-displaying bacteriophage.

1.3. SH3 domains

SH3 domains are prototypes of modular protein domains. In eukaryotic genes the SH3 domain is one of the most abundant modular domain, indicating its essential role for the cell. The human genome encodes approximately 300 SH3 domains present in approximately 220 different human proteins [5]. Several SH3 domains can be found in a single protein. For example, the scaffold protein Tuba contains six SH3 domains and thus, represents the protein containing the highest number of individual SH3 domains found in one single polypeptide [6]. In average, SH3 domains comprise of 60 amino acids, which fold into a compact beta-sandwich, which consists of five major beta-strands. The first three major strands are connected by three loops, namely RT-loop, N-Src-loop and the distal loop, whereas a 3_{10}helix separates the last two strands [7].

1.3.1. Ligands and mode of interaction

Ligands capable of binding to SH3 domains are mainly peptides which can adopt the conformation of a left-handed polyproline-2 helix (PPII2). As the name indicates, prolines are likely to form this type of helix, which is why SH3 ligands often contain prolines in high frequency. The so-called “hydrophobic groove” of the SH3 domain represents the binding region for these ligands, it accommodates several conserved aromatic residues and is flanked by the RT- and N-Src-loop [8]. The affinity and specificity of proline-rich peptides that bind to SH3 domains are mostly moderate- usually with a K_D in the micromolar range, indicating a transient and dynamic mode of interaction. The necessity of this transient interaction character becomes apparent when considering that SH3 domains frequently function as adaptors in highly dynamic cell signalling cascades. Nevertheless, SH3-mediated interactions can also be highly specific and show high affinities in the low nanomolar range. This is mainly accomplished by additional interaction surfaces on the SH3 domain or the ligand, contacts between other domains of the two interaction partners or the interplay between additional members of a multiprotein complex [9].

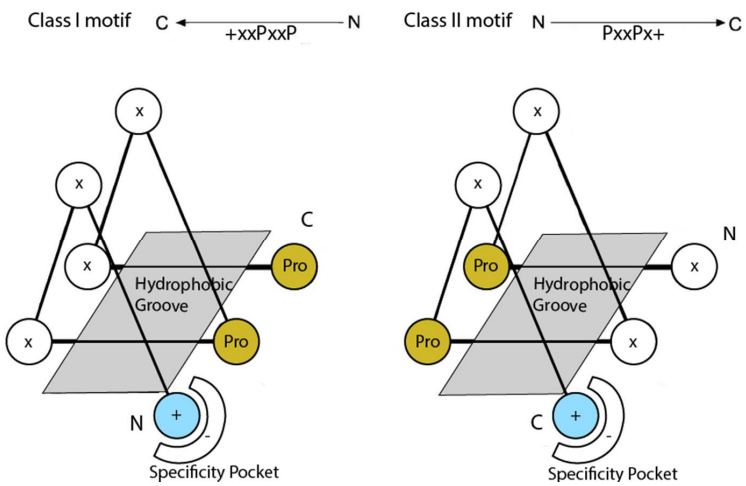

Figure 2. Binding of class I and II ligands to the SH3 domain. Both motifs, class I and II, adopt a PPII2-helix conformation. The essential prolines (dark yellow) and positive amino acids (light blue) of the ligands interact with the hydrophobic groove (grey rhombus) and the specificity pocket of the SH3 domain, respectively. Modified after [7].

1.3.1.1. Typical and atypical binding motifs

The first identified minimal consensus sequence of SH3 ligands is the nowadays commonly known PxxP motif, where “P” represents proline and, “x” stands for any amino acid. The hydrophobic groove of the SH3 domain contains two distinct xP-binding pockets positioning the ligand on the SH3 domain interacting surface. This positioning can happen in two different orientations- depending on the N- or C-terminal location of a positively charged residue, relative to the PxxP motif [7]. The positively charged residue interacts with a negatively charged region of the SH3 domain, the “specificity pocket” [10, 11]. Depending on the ligand’s orientation relative to the SH3 domain binding site, it is classified either as class I ligand (consensus sequence +xxPxxP, where “+” is a positively charged amino acid) or class II ligand (consensus sequence PxxPx+)[9]. Both consensus sequences are understood as “typical binding motifs”, in other words, canonical motifs fitting the binding pockets of the SH3 domain interaction surface. Whereas most of the identified SH3-ligands define as class I or II ligand, a considerable number of SH3 binding partners cannot be grouped into either of the classes, due to their lack of the typical binding motif. Consequently, these so-called "non-consensus ligands” [9, 12] or “atypical motif” containing ligands do not necessarily adopt the PPII-helical conformation or interact with the binding pockets of the SH3 domain hydrophobic groove. Instead, the interaction is often mediated via the “specificity zone”, an area of the SH3 domain surface (including the specificity pocket), which engages with the ligands via extended atomic

contacts. Interestingly, this mode of SH3 domain interaction can result in a remarkably increased affinity and specificity and can outperform the xP-binding pocket in terms of contribution to the overall affinity and selectivity of the interaction [9].

1.3.2. Cellular functions of SH3 domains

SH3 domains are involved in numerous cellular functions. As an integral part of Src-family phosphotyrosine kinases (PTKs), they take part in the regulation of cell proliferation and modification, cell migration, modification of components of the cytoskeleton and cell survival signalling cascades. Substrate recognition and regulation of PTK activity are the main mechanistic functions of SH3 domains in this context [13]. Many SH3 domain-containing proteins, which do not belong to the family of PTKs, are functioning as adaptor molecules in the cell. By binding to their ligand and translocating it to another site or compartment of the cell, for example to the cell membrane (where the input signal originates from a receptor tyrosine kinase [RTK]), the ligand-protein gets into proximity with its substrate or another interaction partner. A well-known example for this scenario is the binding of Growth factor receptor-bound 2 protein (GRB2) SH3 domain to Son of sevenless (Sos) and their translocation to the cell membrane via the interaction of GRB2 SH2 domain with the phosphorylated and thereby active RTK. The increased concentration of Sos at the cell membrane allows the interaction and consequently, the activation of the small G protein Rat sarcoma (RAS) [5]. Adapter molecules which contain several modular domains in addition to an SH3 domain can also function as scaffolding proteins for protein assembly [13].

1.3.3. Diseases involving dysregulation of SH3 domain-mediated interactions

Due to the involvement of SH3 domains in a multitude of cellular functions and signalling pathways, the dysregulation or alteration of SH3-mediated interactions can lead to various pathologies, including cancer. A crucial example of altered SH3 domain-mediated interactions contributing to abnormal cell signalling and potentially resulting in cancer, is the loss of the autoinhibitory function of Abelson (Abl)-SH3 domain in Breakpoint cluster region (BCR)-Abl. Normally engaged in an intramolecular interaction with a proline residue (located between the Abl SH2 and catalytic domain), the SH3 domain contributes to a limited catalytic activity of Abl. Mutations either in the SH3 domain or in its intramolecular contact region can thereby result in a loss of this critical regulatory inhibition of Abl-kinase activity [14, 15]. The ”Disintegrin and metalloproteinase domain-containing protein 15” Adam15 was found to be

associated with prostate and breast cancer progression. Interestingly, Adam15 exists due to alternative splicing, in different isoforms in the cell. Those isoforms contain different sets of proline-rich motifs and bind consequently to different SH3 domain-containing proteins or, like the isoform 1, do not support SH3-mediated interactions at all [16]. The dysregulation of alternative splicing of Adam15 can alter the interaction profile of the protein dramatically and contributes to the development of cancer [17].

1.3.4. SH3 domains employed by pathogens

Considering the ubiquitous occurrence of SH3 domains in the cellular protein landscape and their importance in cellular signalling and physiology, it is apparent why pathogens adopted this domain from their hosts [18] or developed ligands, which can bind cellular SH3 domains by incorporating so-called "short linear binding motifs"[19], into their proteins [20]. The most prominent examples of pathogen-acquired SH3 domains are v-Src, a viral oncogene identified from the avian Rous sarcoma virus [21] and v-Crk, likewise a viral oncogene, but originated from avian CT10 virus [22]. Both viral proteins transform and promote tumour development in chickens [18]. V-Src is a constitutively active tyrosine kinase, constantly triggering cellular signalling pathways, what consequently leads to cancer development [23]. V-Crk, initially named "$p47^{gag\text{-}crk}$" [24] is a fusion of the viral gag-protein and the "CT10 regulator of kinase" aka Crk. Later, the viral fusion protein was identified as being structurally identical to the SH2- and SH3 domain of cellular Crk, indicating its host cellular origin [25].

1.3.4.1. Viral SH3 domain ligands

Instead of acquiring and employing an entire SH3 domain, many pathogens, in fact mostly viruses, express ligands capable of binding to cellular SH3 domains. This so-called "molecular mimicry"[26] allows pathogens effectively to interfere with or hijack host cellular functions and signalling mechanisms. One prominent representative is the HIV-1 Nef protein, which was found to bind cellular Hck and thereby promoting viral replication *in vitro* [27]. The Hepatitis C virus NS5a protein is described to inhibit Bin-1 induced apoptosis by establishing a direct interaction with Bin-1 [28]. Further examples are the promotion of Alphavirus replication by the interaction of viral nsP3 proteins and cellular Amphiphysins [29], the inhibition of the Raf/ MAPK signalling pathway by HIV-1 Tat binding to Grb2 [30] and others [31]. The binding of *Escherichia coli*-secreted protein F-like protein encoded on prophage U ($EspF_U$) to insulin receptor tyrosine kinase

substrate (IRTKS) SH3 domain, is one example of bacterial proteins functioning as ligands for host cellular SH3 domains. This interaction is special, since due to the occurrence of tandem PxxP motifs in $EspF_{U,}$ it is one the most affine SH3 domain interactions found in nature. A consequence of this interaction is the modulation of the cellular actin assembly machinery by *E.coli* [32]. An example of particular interest for this doctoral thesis was reported by Heikkinen and colleagues, who discovered the binding of avian and Spanish Influenza A virus NS1 protein to cellular Crk(L) proteins, which was found to increase the activation of PI3kinase signalling [33].

1.4. Influenza A virus (IAV)

The host range of IAV is remarkably wide. It was isolated from humans, domestic animals like cats, dogs, pigs and horses, from marine mammals as well as wild and domestic birds like ducks, geese, gulls or terns. However, wild birds are believed to serve as the primary natural reservoir of the virus. Transmissions of IAV from wild birds to mammals are not frequent, yet they are repeatedly observed. Furthermore, the transmitted reservoir strains can persist and circulate in their new host [34].

1.4.1. Disease and epidemiology

Worldwide, IAV is one of the most common causative agents of respiratory infections. An infection with IAV typically leads to a sudden onset of high fever, pain in the muscles, headache, a severe feeling of illness and fatigue, but also to sore throat, cough and coryza [35]. Due to its morbidity and the high number of fatalities, IAV poses a substantial burden to global health. The number of deaths, which yearly account to the so-called “seasonal flu” range between 250000 to 500000 [36]. It is assumed, that large outbreaks of IAV, so-called “pandemics”, recurrently happen since the middle ages. Already Hippocrates described a disease with influenza-like symptoms in 412 BC. However, scientists still dispute whether the described symptoms fit better a diphtheria diagnosis [37]. Nevertheless, after 1580, when the first commonly agreed IAV pandemic occurred, numerous pandemics have taken place, costing millions of lives throughout the centuries. The probably most famous and fatal pandemic was the so-called “Spanish flu” from 1918, which alone caused an estimate of 50 million deaths worldwide [38]

1.4.2. Classification, viral architecture and life cycle

Influenza A viruses represent one of seven genera of the *Orthomyxoviridae* family. As typical for orthomyxoviruses, the genome of IAV consists of negative single-strand RNA segments, which are bound to viral nucleoproteins (NP) [34]. The total genome size is estimated to be about 14 kb [39]. The definition of IAV subtypes bases on the viral surface glycoproteins hemagglutinin (HA) and neuraminidase (NA), whereas 16 different HA and nine different NA glycoproteins exist. Interestingly, human epidemics have been caused by only three of the 16 HA types- HA1, -2, and -3 and two types of NA- NA1 and –2. Under the electron microscope, IAV shows a spherical or filamentous shape. The filamentous form often exceeds 300 nm in length, whereas the diameter of the spherical form is about 100 nm. A host cell-derived lipid bilayer surrounds the virion and is spiked with the HA and NA glycoproteins. In addition, Matrix 2 (M2) ion channels cross the viral membrane. Notably, the number of M2 channels embedded in the membrane is with an M2 to HA ratio of 1:10 to 1:100 comparably small. Beneath the membrane, a matrix consisting of M1 protein surrounds the viral core. This inner part of IAV contains the ribonucleoprotein complexes (RNP) which are RNA segments wrapped around NP, along with the RNA-dependent RNA-polymerase heterotrimer (consisting of the subunits PB1, PB2 and PA) and the nuclear export protein (NEP or non-structural protein (NS 2)[34].

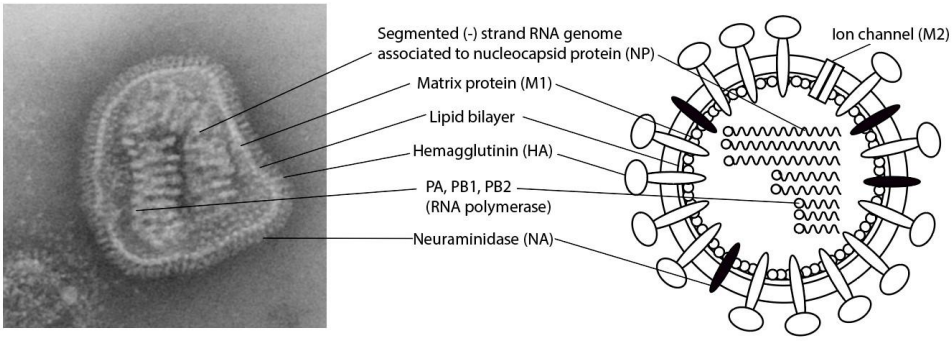

Figure 3. Transmission electron microscope (TEM) image and schematic depiction of an IAV. TEM Credits: CDC/ Dr. Erskine. L. Palmer; Dr. M. L. Martin

To infect the host cell, IAV binds to sialic acid-containing cell surface receptors and enters the cell via endocytosis. In the next step, the viral and endosomal membranes fuse what causes the RNPs to liberate into the cytoplasm. Subsequently, the RNPs are transported to the nucleus, where the viral RNA-dependent RNA-polymerase transcribes the negative-strand viral RNA (vRNA) into positive-strand complementary cRNAs, which serve as templates for vRNA production. The same viral RNA-polymerase also transcribes the vRNA into 5' capped and 3' polyadenylated mRNA, which allows the viral mRNA to be exported and translated by host cell mechanisms. The assembly of new virions happens at the host cell membrane, where newly assembled vRNPs (which are first exported from the nucleus) are, along with the M1 protein, the NEP, and NS1 protein, included in the budding IAV progeny [40].

1.4.3. Genome organization

As mentioned above, the genome of IAV consists of eight single-stranded, negative-sense viral RNA segments. The numerical annotation of the segments follows their size in kb, starting with the largest segment. Table 1 illustrates the eight segments, their translated proteins and protein functions. The segments 7 and 8 are the only ones from which more than one protein is generated by alternative splicing. Table 1 shows the ten essential viral proteins (highlighted in bold), which are translated from the eight IAV RNA segments. Additionally, two proteins are listed, which are expressed from the PB1-segment by alternative start codon usage: PB1-frame 2 (PB1-F2) and PB1-N40 protein. The existence of other accessory proteins, like PA-N182, PA-N155, PA-X, PB2-S1, M42 or NS3 is strain-dependent [41].

Segment	length in nucleotides	Protein/s	length in amino acids	Protein function
1	2341	**Basic polymerase protein 2 (PB2)**	**759**	Polymerase subunit, mRNA cap-binding
2	2341	**Basic polymerase protein 1 (PB1)**	**757**	Polymerase subunit, elongation of RNA and viral RNA synthesis, endonuclease function
		PB1-N40	718	Unknown
		PB1-frame 2 (PB1-F2)	87	Function in apoptosis activation
3	2233	**Acidic polymerase protein (PA)**	**716**	Polymerase subunit with proteolytic activity
4	1778	**Hemagglutinin (HA)**	**550**	Surface glycoprotein, mediates binding to the receptor and fusion, major antigen
5	1565	**Nucleoprotein (NP)**	**498**	RNP component, regulates nuclear import
6	1413	**Neuraminidase (NA)**	**454**	Surface glycoprotein, sialidase activity, release of the virus
7	1027	**Matrix protein 1 (M1)**	**252**	Matrix protein, interacts with the RNP, involved in regulation of nuclear export of RNA and in viral budding
		Matrix protein 2 (M2)	**97**	Ion channel, involved in uncoating of the virus and its assembly
8	890	**Nonstructural protein 1 (NS1)**	**230**	Multifunctional host cell defence antagonist
		Nonstructural protein 2 (NS2/ NEP)	**121**	Nuclear export protein

Table 1. Genome segments encoded by IAV and their corresponding proteins and functions. Proteins in bold letter represent essential viral proteins, modified after [30].

The segmented genome of IAV allows the so-called phenomenon of “antigenic shift”, which can pose a severe threat to the human population by potentially causing new pandemics. The term “antigenic shift” describes a significant change in the hemagglutinin structure of an IAV newly introduced to humans, compared to HA structures already known to the human immune system. This can happen via two different mechanisms- either by the direct transmission of an unknown IAV strain from non-human species to humans or by reassortment events between different IAV strains simultaneously infecting one host cell, resulting in a new IAV strain [42].

1.4.4. Non-structural protein 1 (NS1)

Segment 8, which is the smallest RNA segment of IAV, encodes via alternative splicing for two proteins- the NS1 protein and the nuclear export protein (NEP). NS1 is considered the primary host cell immune response antagonist, with an essential role in suppressing the IFN-α/ β primed antiviral state of the host cell. However, the functions of NS1 are manifold [43].

1.4.4.1. Structural composition

The length of the NS1 protein is IAV strain-specific and varies between 230 and 237 amino acids. Nevertheless, the majority of IAV strains encode an NS1 protein of 230 amino acids, resulting in a polypeptide with a molecular weight of around 26 kDa. NS1 divides into two functional domains, the RNA-binding domain (RBD, amino acids 1-73), located at the N-terminus and the C-terminal effector domain (ED, amino acids 85-230/7) where the last 30 amino acids form a so-called "disordered tail". Both domains are separated by a short flexible linker region, whose sequence can vary substantially among different IAV strains. Also, the disordered tail shows significant differences in amino acid composition depending on the IAV strain. The variability of the linker and the disordered tail contribute substantially to the strain-specific functions of NS1 [44].

As the name implies, the RBD binds different RNA species, whereas no sequence specificity exists. A requirement for RNA binding is, however, the homodimerization of two NS1 molecules. In the full-length protein, dimerization is supported by both, the RBD and effector domain. It has been shown, though that the RBD and ED can also independently dimerize [43]. Apart from stabilizing the homodimerization, the main function of the NS1 ED is mediating interactions with host cell proteins.

1.4.4.2. Functions of NS1 in the viral life cycle

A major role of NS1 is counteracting the host cell interferon (IFN) production and thereby suppressing one of the most important innate immune defence mechanisms of the host cell. Upon pathogen invasion, viral RNA is recognized by pattern recognition receptors (PRRs) as foreign, what initiates the production of Type I IFNs (IFN-α/ β), as a first measure of the cell containing progression and spread of the infection [45]. These soluble cytokines act by upregulating the expression of more than 300 IFN-stimulated antiviral genes (ISGs). This gene upregulation takes place inside the cells from which the cytokines are secreted (autocrine), but also in neighbouring cells (paracrine)[46]. NS1 proteins utilize

two different mechanisms to interfere with the IFN signalling, whereas it is important to note, that not all NS1 proteins are capable of executing these two mechanisms similarly. The first mechanism is the so-called “pre-transcriptional limitation of IFN-β induction” [43], where NS1 targets the Tripartite motif-containing protein 25 (TRIM25) ubiquitin ligase, which is essential for retinoic acid-inducible gene I (RIG-I) and downstream IFN activation. The second mechanism or “post-transcriptional limitation of IFN-β induction” [43], happens by binding of NS1 to the cellular cleavage and polyadenylation specificity factor 30 (CPSF30). This interaction inhibits the processing of all cell host pre-mRNA by CPSF30 and subsequently their translation and expression. This also includes the expression of IFN-β and all ISGs [47]. Another way how NS1 antagonizes the host innate immune response is by interfering with the function of the 2’-5’-oligoadenylate synthetase (OAS) and the dsRNA-dependent serine/threonine protein kinase R (PKR), which are two cytoplasmic proteins promoting antiviral activities of the host cell. While the inhibitory action of NS1 on PKR is the result of a direct interaction between the two proteins [48, 49], the effect of NS1 on OAS is believed to be indirect, possibly by outcompeting OAS over dsRNA binding [48]. Further, NS1 has been found to induce the phosphorylation and subsequent inhibition of c-Jun N-terminal kinase (JNK). This leads to downstream inhibition of the transcription factors ATF-2 and c-Jun and subsequently to the suppression of IFN-β gene expression [50].

Apart from hampering the IFN production, NS1 is interfering with a plethora of other host cellular pathways by interacting with a large number of host cell proteins. One example is the interaction of avian NS1 with PDZ domain-containing proteins, like Scribble and Dlg-1, via a C-terminally located PDZ binding motif. The interaction of NS1 with Scribble was found to protect the IAV infected cells from proapoptotic actions of Scribble [51]. Additionally, both PDZ domain-mediated interactions were found to abrogate cellular tight junctions and thereby increase permeability in IAV infected cells [52]. Another interaction of NS1 with a significant impact on host cell signalling is the one with the regulatory phosphatidyl-inositol-3 kinase (PI3K) subunit p85β. By binding to the inter SH2-domain of p85β, NS1 releases inhibitory contacts between p85β and p110, what results in the activation of PI3K/Akt signalling [53, 54]. With Akt having over more than 100 protein substrates, the PI3K/Akt signalling pathway has an impact on numerous cellular functions, like apoptosis, proliferation, and cytokine production. [43].

1.4.4.3. NS1 as SH3 domain ligand and interaction with Crk adaptor proteins

In 2008 Heikkinen and colleagues [33] identified a perfect class II SH3-binding motif in NS1 proteins from avian IAV strains, which is absent in nearly all human IAV strains. Only four human isolates were found to contain the C-terminally located SH3 domain binding motif, and strikingly, three of these four isolates were zoonotic transmissions. Among them the IAV strain of the highly fatal Spanish flu from 1918 (A/Brevig Mission/1/18/H1N1). Using a phage library displaying a virtually complete collection of 296 human SH3 domains, the N-terminal SH3 domains of CrkII and CrkL could be identified as preferred binders of the NS1 SH3 binding motif. The same study also showed the interaction of NS1 with all full-length Crk proteins from transfected and infected cells. The Crk-family proteins CrkI/II and CrkL represent a special class of regulatory proteins, called adaptor proteins. Their study provided great insight into the understanding of how "modular protein modules assemble into organized protein-protein networks during signal transduction" [48]. Crk adaptor proteins have an integral role in tyrosine kinase signalling, and thereby integrate signals from different origins, e.g. growth factors, extracellular matrix proteins, apoptotic cells and pathogens. There is striking evidence that alterations of Crk(L) functions lead to pathologies, including cancer and pathogen susceptibility [55].

1.4.4.4. Alteration of host cell PI3-Kinase (PI3K) signalling by NS1-Crk interaction

In addition to the NS1-Crk(L) interaction, Heikkinen and colleagues also identified the functional consequence of this assembly. After analysing the activation status of PI3K by quantifying phosphorylated Akt signals from Western blotting experiments (pAkt, Ser473), a significant increase in pAkt levels could be observed when Crk(L) binding competent NS1 proteins were transfected. In contrast, NS1 proteins incapable of binding to Crk(L) caused only a modest increase in PI3K activity [33]. Interestingly, Crk(L) proteins have been previously indicated in PI3K signalling. Reports from Sattler [56] and Gelkop [57] show a direct association of CrkL and CrkII, respectively, with p85, where the binding occurs via a proline-rich motif in p85 and the N-terminal SH3 domain of the Crk proteins. In both studies, the CrkII/L-p85 complex involved a third protein- the Casitas B-lineage Lymphoma E3 ubiquitin-protein ligase (Cbl). The complex plays a role in PI3K signalling upon immune cell activation [57] and signalling in chronic myeloid leukaemia [56].

1.5. Molluscum contagiosum virus

Since the eradication of smallpox in 1980, MCV is the only member of the *Poxviridae* family, commonly establishing infections in humans. The natural reservoir for MCV are humans, whereas diseases resembling an MCV infection have been reported in several other hosts like birds and other mammalian species [58]. A report from Fox and colleagues [59] stated an MCV-like infection in two donkeys. The authors could identify the causative agent as most likely belonging to the Molluscipoxvirus genus (in which MCV is so far the only member). However, none of the isolates could be proven to be an MCV subtype, leaving the known MCV strains restricted to humans. The natural reservoir of MCV is not the only defining feature when comparing it to other poxviruses. Unlike, e.g. variola virus (VAR, the causative agent of smallpox) which systemically infects the human body, the replication of MCV is restricted to the human epidermis. There, MCV causes generally only localized and subacute infections. Remarkably, the virus does not cross the epidermal basement membrane of the *stratum basale* and remains thereby almost invisible to the host cell immune system [60].

1.5.1. Disease and epidemiology

MCV infection can be mainly seen in children (usually 2 to 5 years old), young adults and immunocompromised individuals. A hallmark of the infection is the presentation of defined, glossy, skin-coloured lesions, which have a dome-shaped, umbilicated appearance. The disease can be transmitted directly via skin to skin contact or by sexual contact, whereas the latter leads to an infection of the genital area. Indirect transmission via objects has also been reported [60]. Further, MCV is also considered to be transmitted vertically from mother to child, since several cases of congenital MCV infections are described [61, 62]. The progression and severity of MCV infection can vary dramatically, depending on the immunocompetence status of the infected individual. In healthy, immunocompetent patients, MCV is usually benign and self-limiting. The lesions also called “Mollusca” range from 3 to 5 mm in size. The average duration of infection is 6 to 12 months, although cases have been reported where the infection lasted up to 4 years [63]. This long-lasting persistence of the virus is a consequence of the versatile repertoire of MCV immune evasion strategies [64]. The lesions themselves do not itch, yet, often a localized eczematous dermatitis can be found surrounding the lesions, what causes the sensation of itching to the patients. A frequent consequence is an irritation, secondary bacterial infection and inflammation of the Mollusca, what promotes the further spread of the virus over the skin and causes significant discomfort to the (often very young) patients

[60]. The clinical picture of MCV infection in immunocompromised patients is generally more severe than in humans with a fully functional immune system. Lesions are larger- typically 10- 15mm, therefore the disease is also referred to as “giant molluscum contagiosum”. Furthermore, the number of lesions is highly increased, and they cover large areas of the patient’s skin, including face, neck and scalp [65]. Since the lesions increase in size and number with the progression of immunodeficiency, severe and challenging to treat MCV is often seen in HIV-positive patients with late stages of AIDS and is also recognized as typical comorbidity of the disease [63].

The clinical management of MCV includes physical therapies such as curettage or cryotherapy, the application of chemical agents- e.g. Trichloroacetic acid 100%, Phenol 10%, Cantharidin 0.9% (and others), the treatment with immune modulators like Imiquimod 5% cream or the use of antiviral agents as Cidofovir (either topical or in very severe cases intravenously) [60].

1.5.2. Classification and characteristics

Belonging to the family of *Poxviridae*, MCV is to date the only known member of the Molluscipox genus, comprising four known subtypes- MCV-1, 2, 3 and 4. MCV-1 accounts, with 98% of all cases, for most childhood infections. In HIV-positive patients, MCV-2 is the most commonly seen subtype. Geographical differences in subtype distributions can also be seen, with MCV-3 and MCV-4 being mainly detected in Asia and Australia [63]. Alike other known poxviruses, the genetic information of MCV is encoded on a linear DNA double-strand. The total genome is about 190 kb in size and shows a high content of guanine and cytosine bases- 63%, which is twice as much as the vaccinia virus contains [60]. The viral genome is enclosed in a, for poxviruses typical, dumbbell-shaped protein core, which itself is surrounded by a post-Golgi-derived lipoprotein envelope [64]. The replication of MCV takes place in the cytoplasm of the infected host cell. For that purpose, the virus utilizes a DNA-dependent RNA polymerase, which is included in the virion. Newly assembled viral cores occur initially in the basal epidermal layer of the skin. A hallmark of MCV infection, the viral inclusion bodies also called “Henderson-Paterson bodies” or “molluscum bodies”, appear three to four layers above the basal layer of the epidermis. Recognized as the site in the host cell cytoplasm where the virus is assembled, the “Henderson-Paterson bodies” increase in size, the closer they move towards the granular cell layer. A consequence of the increasing viral inclusion bodies size is the displacement of the host cell nucleus and other organelles. At the final stage of MCV inclusion body maturation, the *stratum corneum* ruptures and releases a high number of infectious viral particles [60].

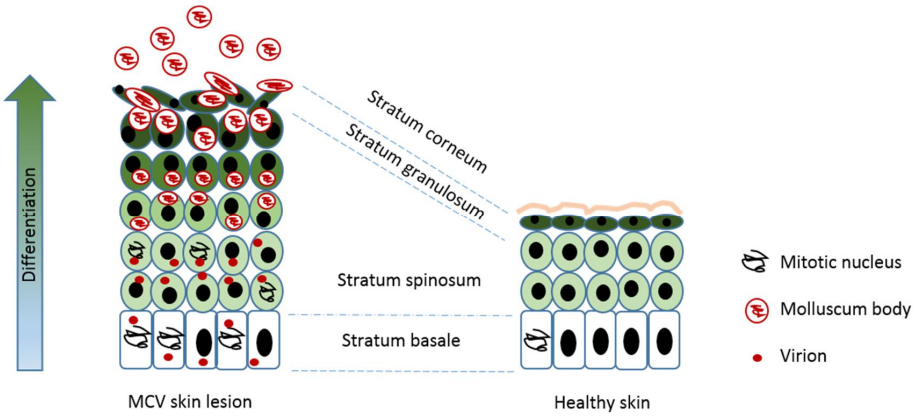

Figure 4. Progression of MCV maturation. Adapted from [53].

1.5.3. Genetic features of MCV

In 1996, the first complete sequence of an MCV genome became available and based on this sequence, 163 proteins were initially predicted, of which 104 were reported to be homologous to vaccinia virus (VV) or VAR proteins [66]. One year later, the number of predicted proteins was corrected to 182 and 105 VV/ VAR homologous proteins [67]. Despite the similarities in the phylogenetic tree of the *Poxviridae* family, MCV is located more closely to parapoxviruses (e.g. Orf virus) than to orthopoxviruses (e.g. VV and VAR) [68]

Like in many other poxviruses, most of the conserved poxvirus genes can also in MCV be found in the central part of its genome. The genes which are exclusive for MCV, are encoded in the flanking region of the viral genome [60]. Since MCV infection is restricted to human keratinocytes, it is most likely that the proteins encoded by MCV-specific genes play a role in virus-host cell interactions. This also includes the versatile immune evasion strategies, which the virus executes via its extensive set of immunomodulatory proteins [64].

1.5.4. MC159 protein

A key player in immune evasion of the host cell by MCV is the MC159 protein. Notably, 77 of the 182 genes of MCV have presumably immunomodulatory functions and only six proteins have been studied in more detail. Remarkably, of those six MCV proteins, MC159 and its close relative MC160 have until this day probably drawn the most scientific attention [64]. Both proteins belong to the so-called FLIP ((FLICE (FADD-like interleukin-1 beta-converting enzyme, aka caspase-8)-inhibitory proteins) family. This important protein family consists of

cellular and viral proteins which regulate cellular apoptosis at the site of the death-inducing signalling complex (DISC) by interacting either with FADD (Fas-associated protein with death domain), Caspase-8 or both [69]. Further, FLIPs have been found to play a crucial role in the regulation of NF-κB and interferon regulatory factor 3 (IRF3) activation [70]. Moreover, Lee and colleagues showed in their work an inhibitory effect of FLIPs on autophagy [71].

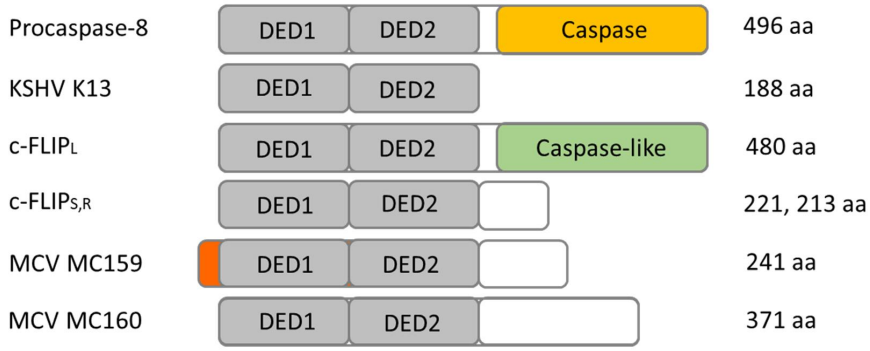

Figure 5. FLIP protein family. In grey: death effector domain (DED); in yellow: caspase domain; in green: caspase-like domain; in orange: α0-helix containing partly PxxP-motif, only present in the DED1 of MC159; white: regions without domain structure adapted from [57]

1.5.4.1. Structural composition

As a member of the FLIP protein family, the 25kDa MC159 contains two tandem death effector domains (DED1 and DED2) which span from position 8 to 175 of the total 241 amino acids (aa) of the protein [64]. Both DEDs of MC159 are clearly distinct from each other in sequence and structure. Furthermore, the DED1 (aa 8-78) differs not only from the neighbouring DED2, also compared to other known DEDs (e.g. FADD DED), it has unique topological features. While DEDs typically comprise six α-helices, the DED1 of MC159 has two additional α-helices-helix H0, located at the very N-terminus and helix H7 and the end of DED1. Furthermore, helix H2 is shorter, and helix H3 is missing in MC159 DED1. Importantly, the crystal structure of MC159 aa 1-187 unveiled a tight association between DED1 and DED2, resulting in a compact, dumbbell-shaped overall structure. The 54 aa long C-terminal tail of MC159 was found to be disordered in the crystal [72], which indicates high flexibility of this region. Thurau and colleagues identified two tumour necrosis factor (TNF) receptor (TNFR) associated factor (TRAF) 3- binding sites in the C-terminal region of MC159 [73]

1.5.4.2. Immunomodulatory functions of MC159

1.5.4.2.1. MC159 and apoptosis

The structural features, defining MC159 as FLIP, implicated early a role of MC159 in death receptor-mediated apoptosis and DISC formation. However, at this moment it was not clear, whether MC159 FLIP has pro- or anti-apoptotic activities. The ability of MC159 to inhibit TNFR- and Fas-induced apoptosis was following the first discovered function of the viral protein [74, 75]. Both DEDs of MC159 are essential for this function and cannot be interchanged or replaced by closely related DEDs [76]. Also, two RxDL motifs located at the C-terminal end of each MC159 DED are crucial for its anti-apoptotic properties [77] Both TRAF binding sites reportedly have a supportive function for the anti-apoptotic activity of MC159 [73]. Several studies addressed the question of how MC159 impairs DISC assembly and thereby prevents death receptor-induced apoptotic signalling [72, 76-78]. Recent work by Fu and colleagues [79] revealed the mechanism by acquiring Cryo-EM structures of Caspase-8 tandem DED filaments, also in the presence of MC159. The so revealed mechanism of MC159 interference with DISC assembly was surprising due to its novel and unique nature. MC159 interacts via its DEDs with FADD and Caspase-8 DEDs and prevents the pro-apoptotic nucleation and polymerization of caspase-8 at the DISC, a mechanism which was named “capping” [79].

1.5.4.2.2. MC159 and the inhibition of TNF-mediated NF-κB activation

Another essential immune modulatory function of MC159 is the protein’s ability to inhibit TNF-induced activation of nuclear factor 'kappa-light-chain-enhancer' of activated B-cells (NF-κB) [81, 82]. Counteraction of NF-κB signalling is in the case of MCV contributing to viral survival since active signalling of this pathway leads to the expression of genes involved in pro-inflammatory and immune responses [83]. Randall and colleagues found that MC159 is interacting with the IκB kinase (IKK) complex [80], which is the master regulator of NF-κB signalling [84]. The same study concluded that the DED1 of MC159 mediates direct interaction with the IKK subunit IKKγ (also known as NEMO) and that this interaction correlates with NF-κB signalling inhibition. Nevertheless, the exact mechanism of how the described MC159-NEMO interaction abrogates NF-κB signalling remained unknown at that time. Just recently, Biswas & Shisler published their work uncovering the underlying mechanism [85]. IKKγ, aka NEMO, needs to undergo polyubiquitination before NF-κB can be activated downstream via IKK. The cellular E3 ligase which ubiquitinates NEMO is the so-called “cellular inhibitor of apoptosis protein 1” (cIAP1). While binding to NEMO, MC159 and cIAP1 were shown to occupy similar surface regions, suggesting a

competitive behaviour of the two proteins towards the binding to NEMO. This hypothesis could be proven correct by the demonstration that MC159 is indeed able to outcompete cIAP1 from binding to NEMO [85].

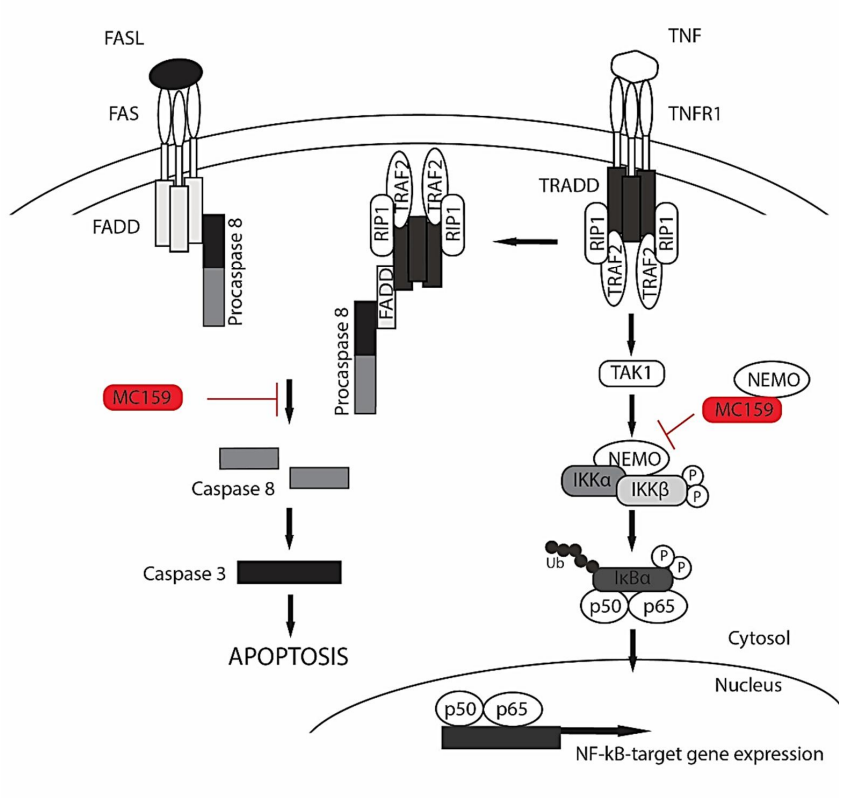

Figure 6. M159-mediated inhibition of death receptor-induced apoptosis and NF-kB signalling. By a novel mechanism, named "capping" [79], MC159 interferes with DISC-assembly and procaspase 8 cleavage, what subsequently leads to apoptosis inhibition. By direct binding to IKK component NEMO and outcompeting cIAP1, MC159 inhibits TNF-mediated NF-κB signalling [80].

1.5.4.2.3. MC159 and IIRF3 inhibition

The activation of interferon response factor 3 (IRF3) is essential for interferon-β gene expression. It is next to apoptosis and NF-κB activation the third component of the "triad of intrinsic immune responses", which is essential for the host cell defence against pathogens [86]. Via interaction with TANK-binding kinase 1 (TBK1) and IKKγ, MC159 prevents TBK1 phosphorylation, which subsequently prevents IRF3 phosphorylation and downstream activation of interferon-β gene

expression. Interestingly, it is the DED2 of MC159, which promotes this function. However, the precise mechanism underlying this observation remains to be elucidated [86].

1.5.4.2.4. Inhibition of autophagy by MC159

In a study from Lee and colleagues from the year 2009 [71], MCV159 was along with other viral, but also cellular FLIPs, identified as a negative regulator of autophagy. By interacting with the Atg3 E2-like enzyme, MC159 and the other FLIPs prevented Atg3 from binding and processing LC3 to LC3-II, which is an essential requirement for autophagosome biogenesis. Moreover, the study identified two critical regions in the viral FLIP's DEDs- the alpha2-helix in DED1 and the alpha4-helix in DED2 (and homologue sites in cellular FLIPs) mediating the interaction with Atg3. Interestingly, the peptides alone were able to bind to Atg3 and FLIPs and could counteract the FLIP mediated suppression of autophagy [71].

1.6. SH3 binding protein 4 (SH3BP4)

In 1999, SH3BP4 was described for the first time after its cDNA was isolated and sequenced from human corneal fibroblasts [87]. The gene encoding the protein is localized on chromosome 2q37.1-q37.2 from which a 5.1 kb messenger RNA is translated. Initially, SH3BP4 mRNA was detected in pancreas, heart, kidney and placenta. However, evidence on protein expression level was missing at that time [87]. Today, 20 years later, SH3BP4 mRNA is known to be expressed almost ubiquitously in the human body, and its expression on a protein level is proven in numerous tissues, including higher expression levels in brain, muscle, lung, liver & gallbladder, pancreas, gastrointestinal tract, male tissues and skin [88].

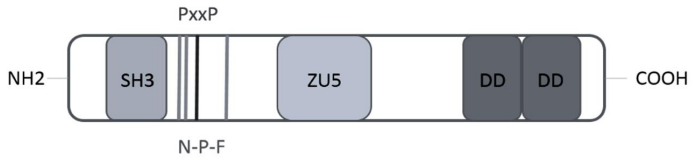

Figure 7. Schematic representation of human SH3BP4. ZU5: domain which is found in mouse ZO-1 tight junction protein and *C. elegans* unc-5 protein; DD: death domain; PxxP motif: SH3 domain binding consensus sequence; N-P-F motif: EH domain binding site, adapted from [82]

1.6.1. Characteristics of protein structure

SH3BP4 is a 963 amino acid polypeptide, which has a resemblance to members of the "Eps15-Homology (EH) network family" [89]. Proteins belonging to this network are mainly involved in endocytosis and vesicle-mediated transport [90]. The structural hallmarks of SH3BP4 support the protein in being an EH-network family member. These hallmarks include three copies of the asparagine-proline-phenylalanine (N-P-F) tripeptide (considered to be the core binding site of EH domains), an SH3 domain, a LIDL-motif (a motif putatively involved in clathrin binding) and a WXXF motif (supposedly mediating interaction with the AP2 adaptor complex) [91]. Further, SH3BP4 contains a C-terminally located death domain (DD) and a ZU5 domain (named after the ZO-1 tight junction protein from mouse and the uncoordinated protein 5 aka unc-5 from *C. elegans*, which both contain this domain). Interestingly, a C-terminal death domain can mostly be seen in eukaryotic proteins containing a ZU5 domain. The same is true for the existence of an N-terminally located SH3 domain. ZU5 domains have been reported to function in ankyrins as spectrin-binding domains, to be involved in apoptosis induction and further in binding to melanoma-associated antigen D1 (MAGE-D1) [92].

1.6.2. Functions of SH3BP4

1.6.2.1. Transferrin receptor-trafficking

The first identified function for SH3BP4 supported the previously drawn, structure-based conclusions for the possible involvement of SH3BP4 in endocytosis and vesicle transport [93]. SH3BP4 or "Transferrin receptor Trafficking Protein (TTP)-how the protein is called in this study, was identified to be selectively involved in transferrin receptor (TfR) internalization regulation, thereby supporting the hypothesis of a "cargo-specific mechanisms of receptor internalization", which the authors proposed. However, the molecular mechanism could not entirely be clarified, leading to three possible scenarios introduced by the study how SH3BP4 aka TTP acts on the regulation of TfR internalization- 1) SH3BP4 functions due to simultaneous binding to TfR and clathrin, with a possible synergetic effect of activating protein2 (AP-2), as TfR-specific adaptor. 2) Function as "scaffold coordinator", where SH3BP4 supports the recruitment of TfR and dynamin (which can also be bound by SH3BP4) to maturing clathrin-coated pits (CCPs) and thereby facilitates the localization of these proteins to the pit/ fission machinery. 3) The so-called "vesicle loading hypothesis", suggests a model where SH3BP4 recruits dynamin via direct interaction to the CCPs and inhibits its GTPase activity. At the same time, the concentration of TfR in the CCPs increases what causes TfR to outcompete

dynamin in binding to SH3BP4. This model would allow time and concentration controlled fission of the CCPs. Importantly, the study also showed that SH3BP4 could be tyrosine-phosphorylated by the epidermal growth factor receptor (EGFR) kinase, which subsequently regulated its interaction dynamin and possibly with other interaction partners [93].

1.6.2.2. FGF receptor trafficking

The involvement of SH3BP4 in fibroblast growth factor receptor (FGFR) signalling identified another cellular pathway were SH3BP4 acts on endocytic and vesicular trafficking [94]. Upon FGFR2b stimulation by FGF-10, FGFR2b gets phosphorylated what leads to the recruitment of PI3K via the SH2 domain of its regulatory subunit p85. P85, in turn, could be shown to associate with SH3BP4, mediated by the SH3 domain of SH3BP4. This interaction is required for sorting of internalized FGFR2b to recycling endosomes and prevents the receptor from being addressed for degradation, which leads subsequently to “increased cell migration and epithelial branching” [94].

1.6.2.3. Negative regulation of mTORC1 signalling

The Rag GTPase complex is a critical component for amino acid-dependent mammalian target of Rapamycin complex 1 (mTORC1) activity. It consists of four members of Rag GTPases- RagA, RagB, RagC and RagD, whereas two heterodimers build up the Rag GTPase complex . Depending on whether GTP or GDP is bound to the RagA/ RagB heterodimer, mTORC1 can be activated, where the GTP bound form mediates mTORC1 activation. The activation of mTORC1 leads subsequently to inhibition of macroautophagy [91]. In this context, SH3PB4 identified as a negative regulator of mTORC1 signalling. By binding and stabilizing the inactive form of the Rag complex under amino acid starvation, SH3BP4 prevents the mTORC1-Rag GTPase interaction. As a result, mTORC1 cannot localize to the lysosome and macroautophagy is induced [95].

1.6.2.4. Additional functions of SH3BP4

A recent study from O'Loughlin and colleagues [96] reported SH3BP4 as a component of the so-called “LIFT complex”, which is a protein complex consisting of the four proteins: myosin 6 (MYO6), GAIP C-terminus-interacting protein (GIPC1), leukaemia-associated RhoGEF (LARG) and SH3BP4. The term

"LIFT" stands for LARG-Induced F-actin for Tethering, and the complex was shown to be essential for modulation of actin at early endosomes [96].

Another novel function of SH3BP4 was proposed by Antas and colleagues [97]. The study stated that by modulating the subcellular localization of β-catenin, SH3BP4 acts through negative feedback inhibition on Wnt-signaling. The ZU5 domain of SH3BP4 could be identified as the region, which is promoting this function [97].

1.7. Cellular autophagy

Autophagy, translated from Greek as "self-eating", defines the cellular process of the highly regulated and coordinated degradation and recycling of cellular and foreign components. More specifically, the degraded components include soluble and aggregated proteins, cell organelles, macromolecular complexes and pathogens [98]. In mammalian cells, three types of autophagy are described, macroautophagy, microautophagy and chaperon-mediated autophagy (CMA). While all three autophagy types share the final step of cargo delivery to the lysosome, followed by cargo-degradation or recycling, the respective upstream processes are distinct [99]. The process of macroautophagy has been extensively studied during the last decades and is in the literature commonly referred to as "autophagy". For this reason and the impact of macroautophagy in the later presented and discussed study, it will be the type of autophagy introduced in the following in more detail.

1.7.1. General overview macroautophagy

Macroautophagy divides into non-selective or "bulk autophagy" and selective autophagy. Non-selective autophagy occurs continuously at basal levels and has a major role in sustaining cellular homeostasis. It is a catabolic process converting cytoplasmic material into metabolites, which can then be used in anabolic and energy-producing processes of the cell [100]. Under conditions of nutrient or energy starvation, the process can further be activated to provide the cell with the necessary metabolites. Unlike bulk autophagy, selective autophagy targets specific objects inside the cell. The terminology used for the different types of selective autophagy depends on the targeted objects. Examples are "mitophagy" (selective degradation of mitochondria), "aggrephagy" (selective degradation of aggregated proteins), "xenophagy" (selective degradation of intracellular pathogens), "peroxypahgy" (selective degradation of peroxisomes) and others [101, 102]. Importantly, each type of selective autophagy is mediated

by a defined set of so-called “autophagy receptors”- proteins which identify their specific cargo and deliver it via an interaction with the autophagy-related protein 8 (ATG8) aka light-chain 3 (LC3) for lysosomal degradation [100]. P62/ Sequestome 1 (SQSTM1) is the autophagy receptor which has been studied intensively, due to its involvement in almost every known type of selective autophagy [103]. Another important autophagy receptor which is involved in xenophagy, aggrepahgy and mitophagy is optineurin (OPT) [104].

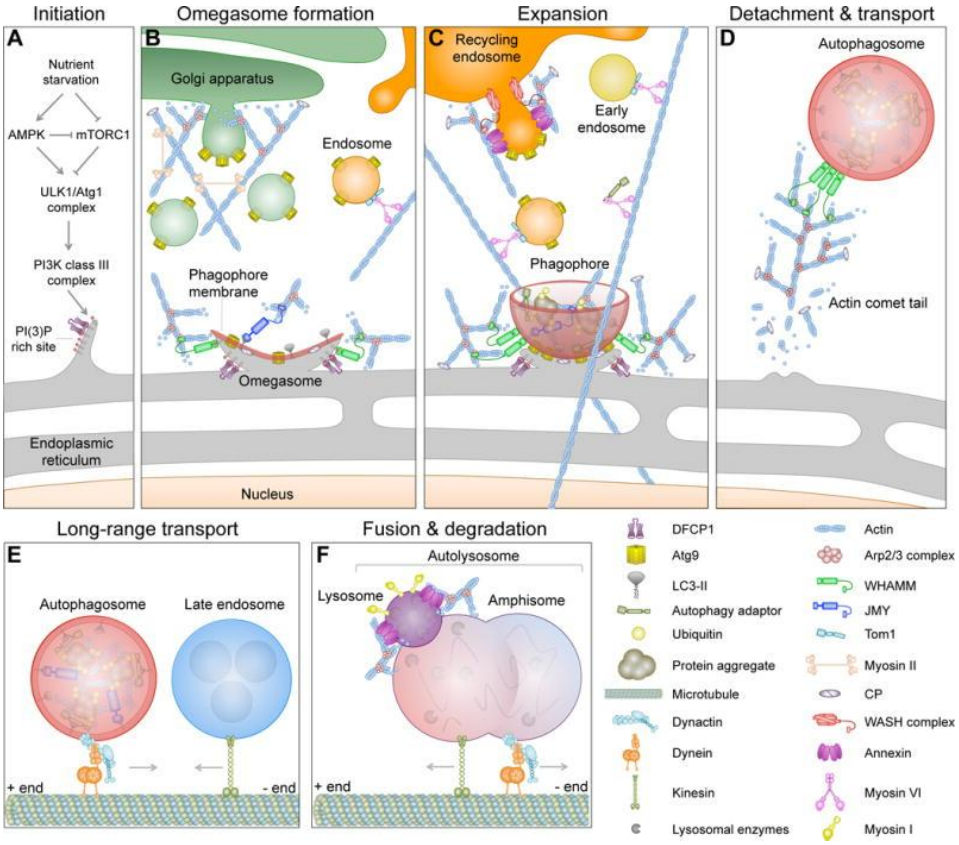

Figure 8. Graphical overview of autophagy stages. © Current Biology, Elsivier, reproduced with permission from Kast et al. [105].

1.7.2. Components and mechanisms

The autophagy process can be divided into six steps promoting the lysosomal elimination of intracellular material of own or foreign origin. When looking at the induction of autophagy, the first step, called “Initiation” (Fig. 8 A) is preceded by signals impinging the cell under starvation conditions. Signals arising from amino acid, energy or growth factor depletion are recognized and forwarded by

different receptors and signalling cascades. However, all those signals lead to mTORC1- the key regulator of mammalian autophagy. When mTORC1 is inactive, autophagy initiation occurs by formation of the ULK-complex and the activation of death-associated protein 1 (DAP1). The regulation of mTORC1 by amino acids is mediated via several amino acid transporters (AAT) and small Rag GTPases. The Rag GTPases form complexes of two heterodimers each containing two of the four different Rag GTPases. This Rag GTPase complex recruits mTORC1 to the lysosomal membrane, which is essential for mTORC1 activity. The regulation of mTORC1 activity by growth factors is, on the other hand, transmitted via the PI3K- Protein Akt/ Tuberous sclerosis (TSC)-Ras homolog enriched in brain (Rheb) pathway. Last but not least, low energy levels lead to 5' AMP-activated protein kinase (AMPK) activation, what in turn inhibits mTORC1 function by activation of TSC2 and phosphorylation of the Regulatory-associated protein of mTOR (Raptor) [105]. Upon initiation, the ULK1/Atg1 complex activates PI3K class III complex, which is recruited to the endoplasmatic reticulum (ER). There the complex generates phosphatidylinositol-(3)-phosphate (PIP3), which leads to sites at the ER enriched in PIP3. These PIP3-rich ER membrane regions are precursors for the forming autophagosomes, what is at this stage, also known as “omegasome” [106, 107]. During the next step, the “formation” (Fig. 8 B), several key proteins are recruited to this forming omegasome and aid mechanically its further enlargement. The growing omegasome extends into the isolation membrane, also known as “phagophore”. The phosphatidylethanolamine (PE) conjugated form of the Atg8 family member LC3- called LC3-II, is inserted in the phagophore membrane. In selective autophagy LC3-II functions as an anchor for autophagy adaptor proteins such as p62 or optineurin, which carry the ubiquitinated cargo to be degraded along with them. As mentioned above, this cargo includes mitochondria, protein aggregates, peroxisomes and intracellular pathogens [103). During the following “expansion” (Fig. 8 C) the phagophore further extends, and more ubiquitinated cargo is delivered. When the phagophore closes, the mature autophagosome will move on to the “detachment and transport” phase, where it releases from the membrane of the ER and is transported first via an “actin comet tail mechanism” [106]. The “long-range transport” (Fig. 8 D) as the next step is supported by dynein-dynactin carrying the autophagosome by moving along microtubules [108]. The last “fusion and degradation” (Fig. 8 F) phase consists of the fusion of the autophagosome with a late endosome, what forms the amphisome and its subsequent fusion with a lysosome, forming the autolysosome. Due to the acidic pH of the lysosome, the pH of the autolysosome is also acidic, which causes the activation of hydrolytic enzymes which subsequently degrade the content of the autolysosome [106].

1.7.3. Exploration of autophagy by viruses

Autophagy is a powerful host cell defence mechanism against viral infections, whereas several different aspects of autophagy promote this function. 1) Autophagosomes trap and deliver viruses which have entered the cell to lysosomal degradation via a form of selective autophagy named “xenophagy” (with the important note that xenophagy traps different types of pathogens, not only viruses) [109]. 2) The autophagy machinery contributes to the innate immune response via pattern recognition receptors (PRR) signalling and subsequent interferon production [110]. 3) Further, it also promotes adaptive immunity by delivering virus-derived antigens for the presentation to T-lymphocytes [111]. During virus evolution, viruses continuously adapt to their host cell by hijacking and modulating cellular signalling pathways to ensure replication and thereby survival. Therefore, it is no surprise that viruses also developed different strategies to oppose or utilize autophagy for their own needs [112]. Depending on the virus lifecycle and characteristics of infection, autophagy is either blocked or employed by different viruses. Persistent viruses like human herpes virus-1 (HSV1), Kaposi’s sarcoma-associated herpesvirus (KSHV), MCV and other DNA-viruses have established numerous methods to avoid or inhibit different steps of autophagy. In contrast to that, RNA viruses like flavi- or picornaviruses depend in certain stages of their lifecycle on the autophagy machinery and exploit it for their needs [112-114].

2. Aims of the study

The aim of this doctoral thesis was the identification and characterization of novel, high-affinity SH3 domain-mediated interactions in the context of i) the human SH3 domain interactome and ii/iii) viral infections, namely, IAV and MCV.

The specific goals were:

i. The identification of selective and strong SH3 domain-mediated interactions from the human SH3 interactome and thereby generate a valuable resource of leads for further investigations

ii. To understand the role and consequences of the SH3 domain-mediated IAV NS1 protein-Crk(L) interaction in the context of PI3K activation and its downstream effector Akt.

iii. The validation of MCV MC159 protein as potential SH3 domain ligand and to identify its preferred SH3 domain-containing interaction partner. Furthermore, to determine the impact of MC159-SH3 binding on host cellular signalling.

3. Material and Methods (study I-III)

3.1. Expression vectors

Expression vectors (I)

To generate the expression vectors for the potential SH3 ligand proteins, 449 open reading frames (ORFs) encoding for these ligands were acquired from the ORFeome v3.1 library (Lamesch, P., Milstein, S. 2007) and inserted via Gateway cloning strategy into the destination vector pEBB/PP-DEST. pEBB/PP-DEST is a modified pEBB/PP vector (gift from Bruce Mayer, University of Connecticut Tanaka et al. 1995) which is driven by an EF-1alpha promotor and contains a 123 amino acid biotin acceptor domain (PP domain)), where a fragment of the modified pDEST vector (gift from Jussi Taipale, University of Helsinki, Finland) has been inserted. The fragment contained the recombination sites attR1 and attR2, the chloramphenicol resistance- and ccdB genes. For individual recombination reactions, the Gateway LR Clonase II Enzyme Mix (Invitrogen, Carlsbad, California, USA) was used according to the manufacturer's protocol. As a result of the Gateway cloning, each potential SH3 ligand ORF was cloned in frame with an N-terminal PP domain.

Expression vectors (II)

The expression vectors for the N-terminal Myc-tagged NS1 proteins (A/Mallard/Netherlands/12/2000/H7N3), A/Udorn/72 and A/WSN/33), were generated by inserting the NS1 cDNAs into pEBB/myc. The cDNA for A/Mallard NS1 generated from total RNA extracted from cells infected with A/Mallard/Netherlands/12/2000/H7N3 IAV, the cDNAs for A/Udorn/72 and A/WSN/33 NS1 proteins were a gift from Ilkka Julkunen (University of Turku, Finland). The C-terminally PP domain tagged Crk proteins (CrkI, CrkII and CrkL) created by inserting Crk protein cDNAs into pEBB/PP. The cDNA for mouse p85β (Open Biosystems) was inserted into pEBB/HA or pEBB/PP for mammalian expression of N-terminal HA-or PP-tagged p85β. All codon changes in NS1 or p85β were carried out by standard overlap PCR mutagenesis. For the generation of the bacterially expressed Glutathione S-transferase (GST)- and maltose-binding protein (MBP) tagged NS1 proteins, NS1 cDNA was cloned into the pGEX-4T-1 vector (GE Healthcare, Chicago, Illinois, United States) and pMAL-c2x (New England Biolabs, Ipswich, Massachusetts, USA).

Expression vectors (III)

The sequence encoding the MC159L gene of molluscum contagiosum virus (EMBL AAC55287.1) was purchased as synthetic gene fragment from GeneScript (NJ, USA) and cloned into two variants of the pEBB mammalian expression vector, containing either an N-terminal hemagglutinin (HA) or PP tag. All amino acid substitutions of the MC159 variants- $MC159^{AxxA(N)}$, $MC159^{AxxA(C)}$ and $MC159^{AxxA(N+C)}$ were generated by overlap PCR mutagenesis. Human SH3BP4 cDNA (NCBI GeneID 23677) was inserted into the mammalian expression vector PCMV/myc-DEST- a modified version of the pCMV-DEST vector, containing a Myc epitope (study **II**) via Gateway cloning. The NF-κB-driven firefly luciferase-expressing vector pBIIX was previously described [115]. To create the lentiviral expression vector pWPI-puro, the EGFP gene in pWPI (from Didier Trono [plasmid 12254; Addgene, Cambridge, Massachusetts, USA]) was replaced with the puromycin resistance gene. For lentiviral transduction and subsequent cell imaging analysis, a fusion construct of each- MC159 and $MC159^{AxxA(N)}$ with a C-terminally added mCherry was subcloned into the pWPI-puro expression vector. For the generation of SH3BP4 Knockout cells used CAG-Cas9-T2A-EGFP-ires-puro expression vector was a gift from Timo Otonkoski (plasmid 78311; Addgene), SH3BP4 genomic RNA (gRNA; BRDN0001487091) was a kind donation from John Doench and David Root (plasmid 78089; Addgene).

3.2. Cell Culture (I, II, III)

Human Embryonic Kidney 293T (HEK293T) cell line from ATCC (I, II, III), human hepatocellular carcinoma (Huh-7, obtained from Mark Harris, University of Leeds, UK) cell line (II), human cervix epitheloid carcinoma (HeLa) cell line (III) and Human Caucasian breast adenocarcinoma (MCF7, stably expressing LC3-EGFP "MCF7/LC3-EGFP", obtained from Marja Jäätelä, DCRC Copenhagen, Denmark) cell line (III) were grown in Dulbeccos Modified Eagle Medium (DMEM) (Sigma Aldrich, St. Louis, MO, USA) supplemented with 4500 mg/L of glucose, 10% fetal bovine serum (FBS) (Gibco, Carlsbad, CA, USA), 0.05 mg/mL penicillin, 0.05 mg/mL streptomycin (Sigma Aldrich), and 1 mM L-Alanyl-L-glutamine (Sigma Aldrich) at 37C in 5% CO2. For the induction of autophagy in MCF7/LC3-EGFP cells, Hank's Balanced Salt Solution with CaCl2 and MgCl2 (HBSS, Gibco, Carlsbad, CA, USA) was applied.

3.3. Generation of knockout cell lines and lentivirus-mediated gene transduction (III)

MCF7/LC3-EGFP were co-transfected with 1.6 µg of SH3BP4 gRNA plasmid (Addgene, Cambridge, Massachusetts, USA) and 0.6 µg of Cas9 expression vector with TransIT-2020 transfection reagent (Mirus Bio, Madison, Wisconsin, USA) following the manufacturer's instructions. 48 hours post-transfection, cells were subjected to puromycin selection for two days at 4 µg/ml puromycin. Afterwards, cells were recovered in regular cell culture medium and subsequently diluted into 96-well plates to perform a single-cell selection.

HEK293T cells were co-transfected with pWPI-MC159-mCherry expression vectors, pDelta8.9 and vesicular stomatitis virus G protein (VSV-G) expressing plasmids using Polyethylimine (PEI) transfection method (3.4.). 48 hours post-transfection, the supernatant was collected and filtered. The filtered lentivirus-containing supernatant was used to infect wild-type or SH3BP4 knockout MCF7/LC3-EGFP cells. Successfully infected cells, stably expressing MC159-mCherry proteins, were selected with puromycin.

3.4. Transfections (I, II, III)

Calcium phosphate precipitation transfection method was applied to HEK293T cells (I, II, III) following the protocol for transfection of a 10 cm cell culture dish (protocol was adjusted according to the number of transfected cells): 20 µg of plasmid DNA were diluted into 250 µl sterile water and then mixed with 250 µl 2,5M CaCl2. The mixture was added dropwise into 500µl 2x HeBS (274 mM NaCl, 10 mM KCl, 1.4 mM NaHPO4, 15 mM glucose, 42 mM Hepes, pH 7.05) under vortexing. The transfection mix was incubated 15- 30 minutes at room temperature (RT) and then added dropwise onto the cell culture medium. The medium and added transfection mix were mixed carefully, and cell culture medium was refreshed 24h post-transfection. PEI (Polyscience Inc., Pennsylvania, USA) transfection was used to transfect HEK293T cells (III) and HeLa cells (III) following the protocol: 12 µg of plasmid DNA were diluted into 300 µl Opti-MEM (Thermo Fisher Scientific, Massachusetts, USA) and mixed with 24 µl PEI. Following an incubation of 15- 30 minutes at RT, the transfection mix was added dropwise in the cell culture dish, which was then chucked carefully to blend the mix into the cell culture medium evenly. Lipofectamine2000 (Invitrogen, Carlsbad, CA, USA) was used according to the manufacturer's instructions for transfection of Huh-7 cells (II). For luciferase assays (III), HEK293T cells were transfected using Fugene6 (Promega, Wisconsin, USA).

3.5. Antibodies and reagents (I, II, III)

Primary antibodies used in (II): mouse anti-Myc (9E10, Santa Cruz Biotechnology, Dallas, Texas, USA), mouse anti-CrkL (clone 5–6, Merck Millipore, Burlington Massachusetts, USA), mouse anti-HA (F-7, Santa Cruz Biotechnology), rabbit anti-phospho Akt(Ser473) (D9E, Cell Signaling Technology, Danvers, Massachusetts, USA), mouse anti-Akt (40D4, Cell Signaling Technology), mouse anti-p85β (T15, AbD Serotec, Kidlington, UK), mouse anti-α-tubulin (DM1A, Sigma-Aldrich, St. Louis, Missouri, USA) and guinea-pig anti-NS1 [116]

Primary antibodies used in (III): mouse anti-Myc (clone 9E10; Santa Cruz Biotechnology, Texas, USA), mouse anti-SH3BP4 (A6; Santa Cruz Biotechnology), mouse anti-caspase-3 (BD Bioscience, California, USA), mouse anti-caspase-8 (1C12; Cell Signaling Technology) rabbit anti-LC3B (D11) XP (Cell Signaling Technology), mouse monoclonal (6G6) to red fluorescent protein (RFP) (ChromoTek, Munich, Germany), rabbit anti-GFP (sc-8334; Santa Cruz Biotechnology), and rabbit anti-p62 (Enzo, New York, USA).

Secondary antibodies used: IRDye 680CW goat anti-mouse IgG (II, III) IRDye 800CW goat anti-mouse IgG (II), IRDye 680CW goat anti-rabbit IgG (II) IRDye 800CW goat anti-rabbit IgG (II, III), Streptavidin IRDye680CW (II), Streptavidin IRDye800CW (I, II, III) were from Li-Cor Biotechnology (Lincoln, Nebraska, USA). IRDye800CW rabbit anti-guinea pig (II) was purchased from Rockland Immunochemicals (Pottstown, Pennsylvania, USA).

Reagents: Imatinib (II, c-Abl inhibitor) was purchased from Sigma-Aldrich. Cycloheximide was a kind gift from Tomas Strandin (University of Helsinki, Finland).

3.6. Recombinant protein expression (I, II)

GST-fusions of SH3 domains (I) and NS1 proteins (II) were expressed in *E.coli* BL21 and affinity purified using glutathione-sepharose 4B (GE Healthcare). MBP-NS1 proteins (II) were expressed similarly and affinity purified with amylose resin (New England Biolabs, Ipswich, Massachusetts, USA).

3.7. Phage Display library-based screening (I, III)

3.7.1. Protein expression and cell lysis (I)

10 cm cell culture dishes of HEK293T cells were transfected with pEBB/PP ORF expression vector or pEBB/PP control plasmid by standard calcium phosphate precipitation method (see “Transfections”). 36 h post-transfection, cells were

collected in ice-cold PBG buffer (1× phosphate-buffered saline (PBS) supplemented with 10% glycerol and 0.5% Tween-20, supplemented with protease inhibitors “complete” from Roche, Basel, Switzerland). Cell lysis was continued by sonication at 0.2- 0.3 kJ on ice using a Bandelin Sonoplus homogenizer (Berlin, Germany). The cell lysates were cleared by centrifugation, 20 minutes at 13200 rpm and 4°C.

3.7.2. Protein expression and cell lysis (III)

Subconfluent HEK293T cells in 10 cm cell culture dishes were transfected with pEBB/PP-MC159 constructs either by standard calcium phosphate precipitation method or using PEI transfection protocol (see 3.3. [III]) 48 h post-transfection, cells were collected in ice-cold TEN (50 mM Tris [pH 7.4], 1 mM EDTA, 150 mM NaCl) buffer. Afterwards, cells were pelleted at 4°C, and 1500 rpm and the cell pellet was frozen at -80°C. Cells were thawed on ice and lysed in a volume of 1 ml lysis buffer (phosphate-buffered saline [PBS], 1.2% [V/V] NP-40, protease inhibitor [Thermo Fisher Scientific, Massachusetts, USA]). Following, the cell lysates were sonicated twice, 10 seconds at 50% amplitude with a Bandelin Sonoplus homogenizer and cell debris was removed by spinning the cell lysates at 4°C and 132000 rpm for 20 minutes.

3.7.3. Protein pull-down and panning

PP-tagged ligand proteins were precipitated with streptavidin-coated magnetic beads (Dynabeads M-280-Streptavidin, Invitrogen, Carlsbad, California, USA) by incubation of the cell lysates with magnetic beads for 90 minutes at 4°C under rotation. Following, the beads were washed one time with 1 ml cold lysis buffer and another two times with 1 ml cold PBST (0.05% Tween-20 in 1× PBS). Each wash was done for 1 minute under rotation.

For phage-library panning, the magnetic beads, containing the ligand-protein, were incubated with a mixture of human SH3 library-displaying phages (about 1 × 1011 colony-forming units [CFU]) and 5% milk in PBST. The ratio of phage solution to milk in PBST was 2:1, and the bead-phage mix was incubated at RT and under rotation for 120 minutes (I) or 90 minutes (III). After incubation, beads were washed with 1 ml PBST four times for 5 minutes under rotation, whereas the reaction tubes were changed after the first and third wash. Meanwhile, TG1 bacteria were grown to log-phase (OD 0.5- 0.6), and 800 µl of TG1 bacteria were mixed with the beads after the last wash, containing now the ligand-protein and the bound phages. The mix was incubated, shaking at 37°C for 1 hour. Finally,

bacteria were plated in different dilutions lysogeny broth (LB) plates containing ampicillin. To identify the ligand-protein interacting SH3 domains, the formed bacterial colonies and their SH3 domain-encoding phagemids (pG8J8/SH3 clones) were subjected to Sanger sequencing.

3.8. Cellular protein interactions (II, III)

3.8.1. Immunoprecipitation and protein pull-down (II)

Transfected HEK293T cells were collected in ice cold TEN buffer as described in (3.6.2.) and lysed in 1% NP40 buffer (150 mM NaCl, 50 mM Tris-HCl, 1% [V/V] NP40, pH 7.9, including protease and phosphatase inhibitors). For immunoprecipitation, cell lysates were incubated with protein G magnetic beads (Dynabeads, Invitrogen, Carlsbad, California, USA), which were beforehand conjugated with anti-CrkL or anti-HA antibodies. For pull-down experiments, cell lysates were incubated with streptavidin-coated magnetic Dynabeads (Invitrogen, Carlsbad, California, USA).

3.8.2. Protein pull-down (III)

HEK293T cells were transfected as described in (3.4.). Importantly, for interaction studies of MC159 and SH3BP4, the ratio of transfected pEBB/PP-MC159 to PCMV/myc-DEST-SH3BP4 plasmid DNA was 3:1. Cells were collected as described in (3.6.2.) and lysed in 1 ml of HEPES-based lysis buffer (40 mM HEPES, 120 mM NaCl, 10 mM glycerophosphate, 1% [V/V] Triton X-100, including protease and phosphatase inhibitor minitablets [Thermo Fisher Scientific, Massachusetts, USA]). Lysates were incubated on ice for 20 min and cleared by centrifugation. Following, the lysates were incubated for 90 min at 4°C, under rotation, with streptavidin-coated magnetic Dynabeads (Invitrogen, Carlsbad, California, USA).

3.9. Gel electrophoresis and Western blotting and detection (I, II, III)

All protein samples subjected to Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were mixed before with 4x Laemmli buffer (200 mM Tris-HCl pH 6.8, 8% w/v SDS. 0.4% w/V bromophenol blue. 40% V/V glycerol, 5% V/V β-mercaptoethanol), boiled 5 min at 95 C and immediately transferred on ice. After preparation, protein samples were separated by SDS-PAGE and transferred on nitrocellulose by semidry Western blotting technique (Biorad,

Hercules, California, USA). Visualization of specific protein signals was done using the Odyssey infrared imaging system from LI-COR Biosciences.

3.10. Peptide Array (I)

The peptides were synthesized by the former "Peptide and Protein Laboratory" (Haartman Institute, University of Helsinki, Finland) and spotted on cellulose using the Multipep system (Intavis Bioanalytical Instruments, Cologne, Germany). The SPOT synthesis was carried out according to the manufacturer's protocol [117]. Parallel arrays were prepared on glass slides with the SlideSpotter equipment (Intavis Bioanalytical Instruments). The spotted peptide arrays were incubated in blocking solution (PBST, 5% w/V nonfat dry milk) for 2 hours and washed three times with PBST. SH3 domains which were identified by the preceded phage screen as strong interaction partners to at least one of the ligand proteins (Lyn, Src, Fyn, Tec, Yes, Crk, CrkL, Amphiphysin (AMPH), CMS (1/3), BAIAP2L1, ArgBP2, Eps8L3, Intersectin 1 (Itsn1) (3/5)) were expressed and affinity-purified from bacterial cultures as GST-fusion proteins. The peptide arrays were incubated with individual SH3 domains, which were beforehand diluted to a final concentration of 1 µM in blocking solution, for 1 hour at room temperature. Following, the slides were washed three times with PBST and once with PBS. To detect SH3 domain binding, peptide array slides were incubated with anti-GST antibody, which was later conjugated with dye IRDye 800CW (LI-COR Biosciences). The signals were imaged and quantified using the Odyssey infrared imaging system from LI-COR Biosciences. The interaction strength of the individual SH3 domains with each peptide was estimated from the detected signal of the corresponding spot in the arrays. The binding was considered specific, if the signal exceeded or was equal to 50% of the highest intensity value from the evaluated array.

3.11. Semi-quantitative pull-down assay (I)

The SH3-GST-pp fusion proteins were expressed and purified as described before (3.5.). The purified fusion-proteins were dialyzed against NP40 lysis buffer (pH 7.4) and adjusted to a final concentration of 0.5 µM. Free biotin was added to the solution to a final concentration of 500 µM. The SH3 domain ligand proteins were expressed and captured by streptavidin-coated magnetic Dynabeads (Invitrogen, Carlsbad, California, USA) from HEK293T cells as described in 3.6.2. and 3.6.3.. To saturate all remaining free biotin-binding sites on the streptavidin-coated magnetic beads after ligand-immobilization, free biotin was added to a final concentration of 500 µM. Following, the beads containing the ligands were

washed three times with NP40 lysis buffer and incubated with the recombinant expressed and purified SH3-GST-pp fusion proteins for 1 hour at 4°C under rotation. The magnetic beads, containing the ligand and putatively bound SH3-GST-pp fusion proteins, were washed three times with lysis buffer. The protein complexes were separated and analysed by SDS-PAGE and Western blotting as described in (3.8). The signals of the bound SH3-GST-pp fusion proteins were quantified with the Image Studio™Lite Software (LI-COR Biosciences) and plotted against the ration of these signals vs the total concentration of SH3-GST-pp fusion protein incubated with the ligand. To calculate KD-values, a linear regression model was applied to the plotted data, and the regression line slope provided the negative KD-value.

3.12. NF-κB luciferase reporter assay (III)

HEK293T cells were transfected with 50 ng of pEBB-HA-RenLuc, 50 ng of NF-κB-driven pBIIX-fLuc DNA and pEBBHA-MC159 DNA using Fugene 6 (Promega, WI, USA) in accordance with the manufacturer's protocol. At 24 h post-transfection, the cell culture medium was changed to medium containing 10 ng/ml TNF-α. After 8 hours of treatment, cells were washed with ice-cold PBS and lysed on the plate in 100 µl passive lysis buffer (Promega, WI, USA). To determine luciferase activities, the dual-luciferase reporter assay system from Promega was used following the manufacturer's protocol. The measured Renilla luciferase was used as an internal control.

3.13. TNF-induced apoptosis (III)

HeLa cells were transfected either with pEBBHA-MC159 DNA or control plasmid. The treatment of cells with 10 ng/ml TNF-α and 10 µg/ml cycloheximide was done 24 hours post-transfection for 8 hours. Following treatment, cells were washed with ice-cold PBS and lysed on plate in 200 µl NP40 lysis buffer (pH 7.4). All samples were analysed as described in 3.9.

3.14. Autophagy analyses (III)

MCF7 cells expressing LC3-EGFP (macroautophagy marker) alone or stably co-expressing wild-type MC159- or MC159$^{AXXA(N)}$-mCherry fusion proteins were seeded on 13-mm-diameter coverslips, 1.5 mm, 24 hours after seeding and two washes with PBS, cells were either incubated with HBSS (starvation) or complete cell culture medium (untreated) for 6 hours. Following, cells were fixed with 4%

paraformaldehyde (PFA) for 15 min and washed three times with PBS. Subsequently, mounting of the cells on microscope slides was done using Hoechst stain (nuclear marker) containing mounting medium. Imaging of the slides was done using an Axio Imager Z2 upright epifluorescence microscope (Carl Zeiss Microscopy, GmbH, Germany) with a 20× (numeric aperture [NA],0.8) air immersion objective with a Hamamatsu Orca Flash 4.0 LT, 4-megapixel monochrome sCMOS camera. Images of three fluorescence channels were taken- nuclear (DAPI), EGFP, and mCherry, whereas three image stacks were taken for each position. The Huygens batch processing application (scientific volume imaging; Huygens Software) was used for deconvolution of the images. Three independent experiments, including all 12 experimental conditions (wild type, mutant, or control/starvation or no treatment/parental MCF-7/LC3-EGFP cells or their SH3BP4 knockout derivatives) were performed. Quantitative autophagosome analysis was performed with Imaris image analysis software 9.2 (Bitplane AG, Zurich, Switzerland). Statistical analysis of the data was done by performing a Mann-Whitney U test using IBM SPSS statistics software.

4. Results

4.1.1. Identification of human SH3 interactome ligands using phage-display library screen approach (I)

With the goal to examine selective and affine SH3 domain-mediated interactions in the human interactome, the Human ORFeome v3.1. cDNA library [118], which contains the genetic information of human 10214 genes, was first screened bioinformatically for potential SH3 domain ligands. The screening was based on the identification of SH3 target motif sites in the potential ligands encoded by the ORFeome library. The first of two applied SH3 target motif prediction strategies was based on the identification of conventional SH3 binding motifs such as PxxP as well as unconventional motifs like PxxDY, Px(P/A)xxR, ϕxRPxR, and PxxxRxxKP ("x" stands for any amino acid and ϕ is a hydrophobic residue). In a second approach, potential ligands were identified based on the existence of regions with a high frequency of prolines or basic amino acid residues. We also included several gene expression products from the library in the second approach, which did not contain an apparent SH3 binding motif, but were known to be involved in cellular signalling and contained at least one modular protein domain such as SH3, SH2 or PDZ. The second screening approach was intended to uncover novel types of SH3 interaction.

When combining the results of the described screening approaches, we identified a total number of 449 sequences from the ORFeome cDNA library, which potentially encoded an SH3 binding ligand (Supplementary Table S1 in **I**). The group of proteins which contained at least one class I or II SH3 binding motif, represented with 302 proteins, the largest group among the total 449 identified potential ligands.

To express the potential SH3 binding ligands transiently in mammalian cells, we subcloned the 449 ORFs into the pEBB/PP-Dest mammalian expression vector by Gateway cloning. The pEBB/PP-Dest vector was driven by elongation factor 1 (EF-1α) promoter and provided a biotin-acceptor- "PP" domain as fusion to the expressed proteins of interest. The so expressed polypeptides were captured with streptavidin-coated magnetic beads from HEK293T cell lysates and subjected to a single round of SH3 phage-display library affinity screen (Fig. 1 in **I**). Of the initial 449 identified genes, 324 could be successfully subcloned and expressed from cells, resulting in proteins matching the predicted size (Table 1 & supplementary Table S1 in **I**). To estimate a sufficient expression/ concentration of the individual PP-tagged proteins in cell lysates, we compared their corresponding Western blotting signals to the signal obtained from a cell

lysate containing PP-PAK2 control protein, which was present in a concentration known to be sufficient for successful phage-display screening.

A single round of phage-display affinity selection was applied to the 324 successfully expressed potential SH3 ligand proteins, which revealed 19 proteins (5.6%) as strong and selective SH3 binders. “Strong” binding was defined as > 20 fold-enrichment of infectious phages associated with the potential ligand compared to control beads (containing PP-domain only). “Selective” binding was attributed when the selected SH3 domains were dominated by a single or few individual SH3 domains. Another group of 25 proteins (7.7%) showed a substantial enrichment between three to 20 fold of phages over control, but with a lower SH3 domain selectivity. The majority of the 324 successfully expressed proteins (86%) exhibited no significant phage-enrichment over control (less than threefold) and/ or a lack of specificity (Table 1 in **I**).

Reviewing the existing literature on SH3-mediated interactions, we found that some of the 19 proteins and their preferred SH3 domains, which we classified based on the phage-display screening as “strong” and “selective” SH3 interactions (Table 2 in **I**), were previously described. One example was the reportedly robust and highly selective interaction of GRB2-related adaptor protein downstream of Shc (GADS aka Mona) and Src-homology 2 domain-containing leukocyte phosphoprotein of 76 kDa (SLP-76 aka LCP2)[119]. Another example was the interaction of Hematopoietic cell-specific Lyn substrate 1 (HCLS1) with Lyn SH3 domain [120]. Apart from these examples, all remaining interactions we identified as strong and highly specific, were novel.

The results from the phage-display screening further supported the concept that a significant fold-enrichment of phages for a certain ligand was associated with high selectivity (Table 2 in II). For example, only a single SH3 domain was selected by Calcium/calmodulin-dependent protein kinase type IV (CAMK4), Insulin receptor tyrosine kinase substrate (IRTKS) and HCLS1, namely Btk, Eps8L3, and Lyn, respectively. The majority (17 out of 19) of the identified strong and selective SH3 domain ligands comprised in > 50% of the selected phage clones a single SH3 domain. In two cases, Embryonal Fyn-associated substrate (EFS) and Arg/Abl-interacting protein (ArgBP2), 50% or more of the bound phages represented only two individual SH3 domains. For seven of the 19 proteins, we performed sub-screenings in which we used independently produced infectious SH3 phage libraries and ligand preparations. The obtained results (Supplementary Table 2 in **I**) were similar to those from the initial phage screenings and thereby confirmed the excellent reproducibility of this approach.

4.1.2. Analysis and binding profiles of the SH3 ligand protein peptides

Since the 19 identified highly selective SH3 ligand-proteins contained up to 12 potential SH3 binding motifs, it was of great interest to understand the individual binding properties of these motifs and their contribution to SH3 domain interactions. For this purpose, a total number of 65 peptides was synthesized. These peptides ranged between 15 and 17 amino acids in length, and their sequences resembled the motifs on which we initially based the screening of the ORFeome library for SH3 ligands (Table 3 in **I**).

To perform the peptide array, all 65 peptides were spotted on glass slides and incubated with 16 individual SH3 domains, which were expressed and purified from bacteria as GST-fusion proteins. The 16 SH3 domains included those, which were preferentially bound by at least one of the 19 ligands from the phage library. Due to insufficient recombinant expression or unspecific binding to the glass slides, the SH3 domains of Btk, GADS and Eps8L3 could not be used as probes in the peptide array. In the case of Eps8L3, we were able to produce the homologous SH3 domain of Eps8L1 as a replacement. The 65 peptides were spotted in two different schemes on the glass slides to prevent any position-related bias, and each probing of the glass slides with the individual 16 SH3 domains was carried out in three independent experiments (Figure 2 in **I**).

To evaluate the peptide-binding properties from the peptide arrays, we determined the relative binding signals of the peptides in the context of each individual SH3 domain probe. For this purpose, the background fluorescence signals were subtracted from the peptide spot signals, and the so obtained raw signals were normalized to the average signal from all peptides.

The so determined individual binding capacities of the 65 peptides for the probed SH3 domains were for simplification categorized into “negative”, “weak”, “intermediate” and “strong” interactions. An interaction was defined as “negative” when the individual peptide spot signal was less than 0.5 fold of the average signal of all peptides in the array. “Weak” was attributed for signals 0.50–0.99 times the average peptide spot signal, “intermediate” for 1.0–1.99 times the average signal and all signals with values > 2 were defined as “strong” interactions (Table 3 in **I**). We found, depending on the probed SH3 domain, 4.6–23.1% (3–15 out of 65) of the peptides acting as “strong” binders. “Intermediate” binders were represented by 6.2–43.1% (4–28 out of 65), “weak“ by 4.6–36.9% (3–24 out of 65) and “negative” binders by 26.2–69.2% (17-45 out of 65). The differences between the peptide array results for the individual SH3 domains can be explained on the one hand by their distinctive selectivity, but on the other hand also by the experimental quality, which depended strongly on the signal to noise ratio of each SH3 domain probe.

According to the acquired results, the binding abilities and preferences of the 65 peptides towards the tested SH3 domains varied greatly. While ten peptides behaved rather indiscriminately and showed “strong” or “intermediate” binding to at least 13 of the 16 probed SH3 domains, many others were highly selective in their binding. 37% of all arrayed peptides bound to three or fewer individual SH3 domains. However, six peptides bound “strongly” to only one or two SH3 domains, while being “negative” for all remaining SH3 probes. For example, the ENSSVVIPPPDYLECLSM peptide from IRTKS, which exclusively bound Eps8L1 SH3 domain with exceptionally high affinity (> 10 times over average peptide signal), showed no interaction with any other probed SH3 domain. The Eps8-family SH3 domains are known to preferentially to bind to PxxDY motif-containing ligands [9]. However, the results of the phage display screen suggested Eps8L3 as a preferred native interaction partner of IRTKS, which also outcompeted Eps8L1. Further, the peptide HPLTRVAPQPPGEDDAPY from Efs and DSLPVAPGRDRPPKQPPT from PHF21B showed selective binding to the second of three SH3 domains (2/3) of ArgBP2 and CMS SH3 (1/3), respectively. Other interaction partners which associated with high affinities were the PDZK8 peptide NYKIRFKPFFPYQTLQGF and CIN85 SH3 (1/3) as well as CIN85 peptide QKPSVPAIPPKKPRPPKT and CrkL and CMS SH3 (1/3). Notably, 14 out of 65 peptides did not exceed the average binding signals of all peptides for any of the probed SH3 domains, and only three peptides exhibited negative binding profiles for all 16 SH3 domains (Table 3 in **I**).

Comparing the results of the phage display screenings with those from the peptide array, we found only a moderate correlation. One crucial fact which needed to be considered when we attempted to compare the results from both approaches, was that most of the ligand proteins contained at least one peptide which exhibited a promiscuous SH3-binding character in the peptide array. In fact, when we removed the six most promiscuous peptides (from ligand proteins ASAP2, CBLC, PDZK8 and Efs), which led the list of strongest SH3 binding peptides, only two peptides remained among the upper 10% of peptides which bound to a certain SH3 domain. Namely the ENSSVVIPPPDYLECLSM peptide from IRTKS, which was the strongest binding peptide for Eps8L1 SH3 and the third-best binding peptide of CrkL SH3 (1/2), QKPSVPAIPPKKPRPPKT from CIN85.

We also found that in several cases, peptides of ligand proteins which were highly selective for certain SH3 domains in the phage display screen did not show high affinities towards the same SH3 domains in the peptide array. An example for this scenario was the selection of Src- and Fyn-SH3 domain by PHF21B in the phage display screen, while the proline-rich peptide of PHF21B clearly bound with high affinity to CMS-SH3 (1/3) and scored negative for Src and Fyn.

Due to the excellent reproducibility and overall binding results of the peptide array, we concluded that the technical quality of the array was good and the obtained results valid. However, presumably, because the approach addressed only binding properties of individual peptides, it seemed unable to faithfully recapitulate the interactions of the corresponding native protein ligands with their favoured SH3 domains.

4.1.3. Semi-quantitative analysis of SH3 ligand-protein interactions

Due to the limited overlap of the identified strong and selective SH3 interactions from the phage display screening with the obtained results for the same ligands from the peptide array, we aimed to study these interactions also in solution. This approach was further intended to explore our hypothesis that the interactions, which we identified by phage display screen, were exceptionally robust.

For this purpose, we developed a semi-quantitative protein pull-down assay in which we immobilized PP-tagged ligand proteins, transiently expressed from HEK293T cells. The ligand-coated beads were incubated with solutions of recombinant expressed and purified SH3 domains at a concentration of 500 nM, or for Scatchard analysis, with concentration series of the same protein solutions. The validity and applicability of this semi-quantitative assay were tested by applying it on the well-characterized interaction between Hck SH3 domain and HIV-1 Nef. The K_D-value, which was previously determined for this interaction using surface plasmon resonance (SPR), was 250nM [121, 122]. The Scatchard analysis for the obtained data from the Hck-SH3-Nef semi-quantitative protein pull-down assay resulted in a K_D-value of 189 nM. Since both K_D-values were very close, we concluded that our semi-quantitative assay was, despite its technical limitations, able to produce reliable results for comparably strong SH3-mediated interactions.

The ligand proteins SH3YL1, SH2B, DENND1A, and PHF21B, were chosen to be studied in the semi-quantitative protein pull-down assay, together with the SH3 domains of Src, Lyn, CrkL, and Fyn. After incubation of the immobilized ligand proteins with the recombinant expressed and purified SH3 domains, unbound proteins were washed off, and the amounts of ligand-associated SH3 domains were analysed by Western blotting. When comparing the association of the four SH3 domains to SH3YL1 in the pull-down assay, it was evident that Src-SH3 domain is the preferred target SH3 domain of this ligand. In contrast, the remaining SH3 domains of Lyn, CrkL and Fyn showed only weak interaction with SH3YL1 (Figure 3B in I). These results were in line with phage display screening

results where SH3YL1 bound Src-SH3 domain preferentially (69%). The ligand-protein SH2B selected Lyn-SH3 exclusively as preferred interaction partner in the phage display screen. This finding was also true when we examined SH2B in the pull-down assay, where it preferentially bound to Lyn-SH3 domain and to a lesser extent the remaining SH3 domains (Figure 3B in **I**). DENND1A showed in the pull-down assay a less striking preference towards the four tested SH3 domains. However, the association of CrkL-SH3 domain was slightly increased compared to Lyn- and Fyn SH3 domain and even stronger elevated when compared to Src-SH3 domain (Figure 3D in **I**). Similarly, DENND1A selected CrkL-SH3 domain as the preferred binder (71%) in the phage display screen. It is important to note that none of the DENND1A peptides showed even intermediate binding towards CrkL-SH3 domain in the peptide array, underlining the striking overlap of the pull-down assay with the phage display results for native DENND1A. Inline, also PHF21B preferentially associated with the same SH3 domains in the pull-down assay as in the phage display screen, namely Src- and Fyn-SH3 (Src- and Fyn-SH3 shared 43% of selected clones, Figure 3E in **I**). In summary, the results we obtained from the pull-down assay and the phage display screen, in which we used the full-length, native ligands, showed a good correlation. In contrast, when the pull-down assay results were compared to the peptide array data, only very few overlap could be seen, and in some cases (e.g. PHF21B) the preferred SH3 domains of the isolated peptides were completely different from those selected by the native ligands.

To investigate the affinities underlying the strong and selective SH3 domain-ligand interactions, which were identified by phage display and further supported by the pull-down assay, we applied the same type of semi-quantitative pull-down assay for selected interaction pairs as described for the estimation of the K_D-value of Hck-Nef interaction. The first interaction pair we studied, was HCLS1 and Lyn-SH3 domain. In the phage display screen, Lyn-SH3 domain was exclusively selected by HCLS1, implying a high affinity for this interaction. In fact, when we quantified the Western blotting signals of HCLS1-associated Lyn-SH3 domain from incubations with differently concentrated solutions of the SH3 domain and examined them by Scatchard analysis, a submicromolar K_D-value of 746 nM was estimated. Based on the results from the pull-down assay, the interaction between SH3YL1 and Src-SH3 domain appeared to be of especially robust nature (Figure 3B in **I**) and was thereby of particular interest for evaluation using this approach. Quantification and analysis of the Src-SH3 domain signals, resulting from the semi-quantitative pull-down assay, proved the hypothesized strong affinity of this interaction by revealing a K_D-value estimate of 187 nM. In contrast, when we examined with the same method the interaction between HCLS1 and Fyn-SH3 (which showed either little or no

association in the pull-down assay and phage-display screen, respectively), a K_D-value of 7 µM was observed- indicating a significantly reduced affinity for this interaction.

In summary, the data from this semi-quantitative pull-down assay recapitulated and thereby supported convincingly the results we retrieved from the phage-display affinity screens. Furthermore, this correlation promotes the phage-display affinity screen as a reliable tool to uncover novel high-affinity SH3 domain-mediated interactions. Also, the results implicate that screening approaches using short linear peptides were unable to recapitulate such specific and strong interactions.

4.2. Reorganization of PI3K signalling complex by Crk(L)-binding competent NS1 proteins (II)

4.2.1. Binding of Crk adaptor family proteins to NS1

As identified in the previous study by Heikkinen and colleagues [33], IAV NS1 proteins which contain a C-terminally located class II SH3-binding motif, interact with cellular CrkII and CrkL proteins and as a consequence potentiate PI3K signalling. CrkII and CrkL are encoded by two different genes. The gene from which CrkII is expressed encodes additionally for a shorter form of the protein-CrkI, which is generated by alternative RNA splicing [55]. This 28 kDa long polypeptide contains, similar to CrkII and CrkL, the N-terminal SH2 and SH3 domain but lacks the second, C-terminal SH3 domain (Fig. 1A in **II**). To investigate whether also CrkI binds to SH3 binding competent NS1 proteins and how strong this potential interaction is compared to the already known ones of CrkII and CrkL and NS1, we co-transfected PP-tagged Crk proteins together with the SH3 binding competent NS1 protein from IVA strain A/Mallard/Netherlands/12/2000/H7N3 (from here on referred as “Mallard NS1”) and performed protein pull-down experiments. To confirm an SH3 binding mediated interaction, we extended the experiment by also including SH3 binding deficient NS1 proteins, in which the critical prolines P212 and P215 were substituted by alanines “NS1-AxxA”. The analysis of the pull-down experiments by SDS-PAGE and Western blotting showed clearly a comparable association between all three Crk proteins and Mallard NS1. The absence of co-precipitated Mallard NS1-AxxA protein confirmed the SH3 domain-mediated binding as the driving force of this interaction (Fig. 1B in **II**).

4.2.2. Individual contributions of NS1-p85β and NS1-Crk(L) interactions to activation of PI3K signalling

It was known that both interactions- NS1-p85β [123] and NS1-Crk(L) [33] promote PI3K activation, yet, the question of individual contributions of the interactions and their possible mechanistic relation regarding PI3K activation remained. Hale and colleagues identified a tyrosine at position 89 in NS1 as critical for binding to the PI3K regulatory subunit p85β [123]. As reported in [33] and mentioned above, the enhanced PI3K activation by avian and some human IAV NS1 proteins depends on a functional Crk(L)-SH3 binding motif. To answer the question of relative impact and a mutual functional relation of the two binding sites of NS1 in PI3K activation, we analysed phosphorylation levels of the PI3K substrate Akt in cells transiently expressing either control plasmid only, wildtype Mallard NS1, the Crk(L) binding deficient Mallard NS1-AxxA and the p85β binding deficient Mallard NS1-Y89F (tyrosine 89 was substituted by phenylalanine). In line with previous results from Heikkinen and colleagues, the obtained phospho-Akt levels in cells expressing NS1-AxxA clearly showed a reduction in PI3K activation compared to wildtype NS1 expressing cells. However, only the NS1-Y89F variant was entirely incapable of activating PI3K signalling (Fig. 2A in II), promoting the conclusion that NS1-Crk(L) interaction enhances PI3K activation, but this event was entirely dependent on the NS1-p85β interaction.

4.2.3. NS1 reorganizes the host cell Crk(L)-PI3K complex

SH3 binding-competent NS1 proteins outcompete p85β from binding to Crk(L)

In addition to the already described NS1-Crk(L) and NS1- p85β interactions, Crk proteins have also been reported to interact with p85 [56, 124]. Interestingly, Crk proteins bind via the same N-terminal SH3 domain to NS1 and p85. To investigate the effect of NS1 proteins with different Crk(L)/p85β binding capacities on a potentially pre-existing Crk(L)-p85β complex, we performed immunoprecipitation experiments from either untransfected cells or cells transiently expressing the different NS1 variants. When precipitating endogenous CrkL from cells, p85β was readily co-precipitated when NS1 was absent. The presence of either wildtype Mallard NS1 (Crk[L] binding competent) or Mallard NS1-AxxA (Crk[L] binding deficient) did not show substantial effects on the Crk(L)-p85β interaction. Strikingly, when the p85β-binding deficient, but Crk(L)-binding competent Mallard NS1-Y89F was expressed, a substantial decrease in Crk(L)-p85β association could be detected (Fig. 2B in **II**). Considering that Crk(L) binds p85 and NS1 via its N-terminal SH3 domain, the result indicated

that NS1 outcompeted p85β from binding to Crk(L), most likely as a result of a higher affinity of NS1 PxxP motif towards the Crk(L) SH3 domain.

To prove the hypothesis that NS1 can outcompete p85β from binding to Crk(L), recombinantly expressed and purified GST-NS1-Y89F protein was added in a series of increasing amounts to lysates of cells transiently co-expressing CrkL and p85β. The association of CrkL and p85β was evaluated by Western blotting signals of CrkL-co-immunoprecipitated p85β. The obtained results clearly showed that increasing amounts of added GST-NS1-Y89F displaced p85β from binding to CrkL in a dose-dependent manner. Strikingly, when the CrkL binding-deficient variant GST-NS1-AxxA was used in the same experimental setting, no changes in p85β-CrkL association could be observed (Fig. 2C in **II**). In an additional experiment where an NS1 binding-deficient variant of p85β (V573M) and wildtype GST-NS1 were used, we acquired similar results to those obtained for wildtype p85β and GST-NS1-Y89F, further supporting the initial hypothesis (Supplementary fig. 1 in **II**).

Formation of a trimeric complex consisting of NS1, Crk(L) and p85β

Combining the results from the earlier described co-precipitation experiments (Fig. 1B, 2B & 2C in **II**) with pre-existing information on observed interactions between the studied proteins NS1, Crk(L) and p85β, the formation of a trimeric complex upon NS1 expression involving all three proteins, became evident. However, the question remained, why p85β was detected in CrkL-immunoprecipitates, regardless of whether NS1 was capable of binding CrkL and thereby displacing p85β. A possible answer to this question was the formation of two distinct trimeric complexes, whose architecture depend on the SH3 binding capacity of NS1 (Fig. 3B in **II**). Due to the ability of wildtype NS1 to outcompete p85β from binding to CrkL and its direct interaction with p85β, the association of p85β to CrkL in the presence of wildtype NS1 was most likely indirect, with NS1 functioning as bridging factor (Fig. 3B in **II**, "Complex I"). In contrast, the observed association of p85β with CrkL in the presence of NS1-AxxA (Fig. 2B in **II**) was presumably direct, since NS1-AxxA was incapable of binding CrkL. We hypothesized that in this scenario a second type of trimeric complex (Fig. 3B in **II**, "Complex II") formed, where p85β was instead of NS1 the bridging factor. In that case, we needed to answer the question of why we were previously unable to detect NS1-AxxA when immunoprecipitating CrkL. We considered either the moderate affinity of p85β-CrkL interaction or the comparably limited availability of endogenous p85β as explanations for NS1-AxxA being undetectable from a presumably CrkL-immunoprecipitated "Complex II". Either way, if our assumptions were correct, co-expression of NS1-

AxxA together with increasing amounts of p85β should facilitate the association of CrkL and NS1-AxxA, despite the incapacity of NS1-AxxA binding CrkL directly. As predicted, the amounts of NS1-AxxA co-immunoprecipitated with CrkL increased with the amount of co-transfected p85β in a dose-dependent manner (Fig. 4A in **II**). In contrast, the ectopic expression of p85β did no effect CrkL-wildtype NS1 interaction, which was already detectable when only endogenous p85β was present.

To extend these findings on naturally occurring SH3 binding-deficient NS1 proteins, the same experimental setup was applied using NS1 proteins from influenza strains A/Udorn/72 and A/WSN/33. In line with the previous results, also the amount of naturally SH3 binding-deficient NS1 proteins increased in CrkL-immunoprecipitates with increasing amounts of co-expressed p85β (Fig. 4B in **II**). Notably, when a Crk(L)-binding deficient variant of p85β (P294A, P297A) was co-expressed instead, only wildtype NS1 could be detected from CrkL-immunoprecipitates, but none of the Crk(L)-binding deficient NS1 proteins (Fig. 4B in **II**, right panel). Conclusively, the presented data strongly supported the existence of two distinct trimeric Crk(L)-NS1-p85β complexes, which formed either in the presence of SH3-binding competent (Complex I) or SH3-binding deficient NS1 proteins (Complex II). The recruitment of CrkL to "Complex I" via NS1 seemed to be more effective than to "Complex II" via p85β.

SH3 binding capacity of NS1 determines the architecture of the trimeric NS1-Crk(L)-p85β complex

To investigate the formation of the two trimeric complexes in more detail, interaction studies were performed from cells transiently expressing CrkL, Mallard NS1 and p85β or their binding-deficient variants (NS1-AxxA, NS1-Y89F, p85β (P294A, P297A) and p85β (V573M) in all applicable combinations. The putatively formed complexes were captured by either immunoprecipitation of CrkL or p85β. The results obtained for the CrkL-immunoprecipitated complexes when co-expressing p85β wildtype together with a control plasmid, NS1, NS1-AxxA or NS1-Y89F were as expected similar to those presented in Fig. 2B in **II**. P85β wildtype could be readily detected from the CrkL-immunoprecipitates of all samples in comparable amounts, except in the presence of NS1-Y89F, which dramatically reduced p85β association. All transfected NS1 proteins were present in the analysed immunoprecipitates (Fig. 5 panel **II**, lanes 2 to 5). When the Crk(L) binding-deficient p85β variant (P294A, P297A) was co-expressed, only the NS1 proteins which directly bound to Crk proteins, wildtype NS1 and NS1-Y89F, could be co-precipitated. As expected, p85β variant (P294A, P297A) was only detected from complexes containing wildtype NS1, which in this case

mediated the interaction. NS1-AxxA was absent from the CrkL-immunoprecipitates when p85β (P294A, P297A) was co-expressed (Fig. 5 panel II & III, lanes 7 to 9 in **II**). The ability of p85β in binding NS1 (V573M) had, as predicted, no effect on the CrkL-p85β (V573M) interaction when either no, or the SH3 binding-deficient NS1-AxxA was co-expressed. Since wildtype- and the Y89F variant of NS1 were able to outcompete p85β (V573M) from CrkL-binding, but unable to directly bind it, a substantial reduction of p85β (V573M) co-immunoprecipitated with CrkL was observed (Fig. 5 panel **II**, lanes 10 to 13). Strikingly, when p85β and p85β (P294A, P297A) were precipitated in an identical experimental setting, only wildtype NS1, which was capable of binding CrkL and p85β simultaneously, was causing an increased association of CrkL and p85β. However, when the NS1 binding-deficient p85β (V573M) was precipitated, no increase of associated CrkL could be detected, since NS1 could not recruit p85β (V573M) to "Complex I" (Fig. 5 panel VI, lanes 3, 7 & 10).

Taken together, the presented results strongly supported our hypothesis of two trimeric complexes, each consisting of Crk(L), p85β and NS1, where either in the presence of SH3 binding-competent NS1 proteins “Complex I” with NS1 as bridging factor was assembled or, in the case of SH3 binding-deficient NS1 protein expression, “Complex II” with p85β as connecting element formed.

To rule out that the trimeric complex assembly depended on additional cellular proteins, wildtype p85β and p85β (P294A, P297A) variant as well as wildtype NS1 and NS1-AxxA and a fragment of CrkL, containing the N-terminal SH3 domain, were recombinantly expressed and purified. After combining the proteins, p85β or p85β (P294A, P297A) were precipitated and analysed for co-precipitated proteins. All three proteins could be detected from the immunoprecipitates when wildtype NS1 was co-expressed, allowing the formation of “Complex I” with NS1 functioning as bridging factor (Supplementary Fig. 2A, lane 2 & 8 in **II**). Also, the formation of “Complex II”, where p85β was the connecting partner, was observed when NS1-AxxA was co-expressed with wildtype p85β and the CrkL fragment. In line with the previous results, the formation of “Complex II” was not possible when the CrkL binding deficient p85β (P294A, P297A) was co-expressed (Supplementary Fig. 2A, lane 5 & 10 in **II**).

4.2.4. Effect of the trimeric NS1-Crk(L)-p85β complexes on PI3K signalling

Assuming that enhanced PI3K signalling caused by SH3 binding-competent NS1 proteins is due to increased efficacy in recruiting Crk(L) proteins to NS1-p85β complexes, promoting the intracellular availability of Crk(L) proteins should enhance PI3K activation in the presence of SH3 binding-deficient NS1 proteins.

Indeed, ectopic expression of CrkL increased phospho-Akt levels when NS1-AxxA or A/Udorn NS1 was co-expressed but had only a minor effect in the presence of wildtype NS1 (Fig. 6A &B in **II**).

Due to the involvement of c-Abl in NS1-SH3 binding [125] and p85β phosphorylation [126], the role of c-Abl in wildtype NS1 mediated activation of PI3K was examined by inhibiting c-Abl with Imatinib. No changes in PI3K activation could be observed (Supplementary Fig. 3 in **II**), which led to the conclusion that c-Abl was not involved in the discussed signalling activation.

4.3. MCV MC159 suppresses cellular autophagy via SH3 domain-mediated binding to SH3BP4 (III)

Similar to avian and Spanish flu influenza A NS1 proteins, MC159 protein of MCV was identified as potential SH3 domain interaction partner by scanning viral proteins from the Swiss-Prot/TrEMBL database with the ScanProsite search tool. Using this method, we discovered two consensus SH3-binding motifs in MC159. A class I consensus motif located in the N-terminus and a class II consensus motif embedded in the C-terminal proximity of the protein. As introduced earlier, MC159 belongs to the FLIP protein family and thereby contains two DEDs. The identified N-terminal class I SH3 binding motif overlaps partly with the a0- and the a1-helix of the DED1. In contrast, the class II motif locates outside of the DED-region, but proximal to two TRAF binding motifs (Fig. 1 in **III**).

4.3.1. Phage display on MC159 and identification of cellular SH3 binding partners

To investigate the SH3 binding capacity of MC159 and to identify potential SH3 binding partners, MC159 was expressed and purified from mammalian 293T cells as PP-fusion protein and immobilized to streptavidin-coated magnetic beads. We subjected immobilized MC159 as well as a PP-tagged control to affinity screenings using our virtually complete human SH3 phage display library [127]. Already after a single round of affinity selection, MC159 conjugated beads showed a substantial enrichment of phages (> 20 fold) compared to control beads (Fig. 2A in **III**), suggesting MC159 as a potent SH3 ligand. Sequencing analysis of the MC159 selected phages, identified the SH3 domain of human SH3 binding protein 4 (SH3BP4) as the highly selected binder. In three individual phage library screening experiments, SH3BP4 SH3 domain represented 80 to 100% of the sequenced phage clones (Fig. 2 in **III**).

To investigate the impact of the two consensus SH3 binding motifs of MC159 on SH3BP4 SH3 domain binding, we generated different variants of MC159, being defective either in one or both motifs. For this purpose, both critical prolines of class I and II motifs were substituted by alanines using site-directed mutagenesis (PxxP to AxxA) (Fig. 1 in **III**). The so generated MC159 variants were tested as affinity ligands in SH3 phage display screenings, using an identical experimental set-up. While functional abrogation of the C-terminal PxxP motif [$MC159^{AXXA(C)}$] had no effect on SH3 phage enrichment, disruption of the N-terminal SH3-binding motif alone [$MC159^{AXXA(N)}$] or together with a non-functional C-terminal PxxP motif [$MC159AXXA^{(N+C)}$], efficiently inhibited the enrichment of phages over control (Fig. 2 in **III**). Consequently, we concluded that the N-terminal class I SH3-binding motif of MC159 promotes the binding to SH3BP4-SH3 domain. In contrast, the C-terminal motif could not be demonstrated as a functional SH3 binding site, despite resembling a perfect class II consensus motif.

Since the SH3 domain of SH3BP4 was dominating the affinity selection of the class I motif profoundly, we considered the existence of other SH3 domains binding to MC159, which were possibly masked by the avid binding of SH3BP4 SH3 domain to MC159. Therefore, we generated a phage library which was depleted of the SH3BP4 SH3 domain and repeated the phage affinity screen. However, no enrichment of wildtype MC159 over control beads was observed when using the depleted phage library, indicating SH3BP4 as unique selective MC159 binder among human SH3 domains. Conclusively, the results from the phage screen experiments indicate the SH3 domain of SH3BP4 as the only preferred binder of the MC159 class I SH3 binding motif.

4.3.2. Association of full-length SH3BP4 to MC159 in cells

To validate and extend our findings from the phage display experiments, we transiently expressed a Myc-tagged version of the full-length SH3BP4 together with PP-tagged MC159 proteins in mammalian 293T cells. Streptavidin-coated magnetic beads were used to precipitate PP-MC159 protein complexes, which were subsequently separated and analysed by SDS-PAGE and Western blotting. Full-length SH3BP4 could be readily co-precipitated with MC159 proteins with an intact N-terminal PxxP motif (Fig. 3A, lanes 2 & 4 in **III**). In line with the results obtained from the phage screening experiments, disrupting the N-terminal PxxP motif abrogated the interaction almost entirely, while mutation of the C-terminal PxxP motif had no effect on SH3BP4 co-precipitating with MC159 (Fig. 3A, lanes 3 & 5 in **III**). Notably, a residual binding of SH3BP4 to MC159 was observed when precipitating $MC159^{AXXA(N)}$ or $MC159^{AXXA(N+C)}$ (Fig. 3A, lanes 2 & 4 in **III**). Since the mock-protein pull-down was completely devoid of SH3BP4, the

residual association of SH3BP4 to MC159 suggested additional weak interactions between the proteins.

To exclude that the observed interaction between the MC159 proteins and SH3BP4 was solely a result of protein over-expression, we repeated the experiment expressing only MC159 proteins transiently. Analysis of the precipitated protein complexes showed a similar interaction pattern of MC159 proteins with endogenous SH3BP4.

4.3.3. Impact of SH3BP4 binding capacity on cellular functions of MC159

The inhibition of FAS and TNFR1 death receptor-mediated NF-κB activation and apoptosis induction are well-established functions of MC159 [64, 74, 75]. We, therefore, investigated the role of MC159-SH3BP4 binding in the context of these functions.

To investigate whether binding of MC159 to SH3BP4 had an impact on TNF-mediated NF-κB activation, we tested wildtype and SH3-binding deficient MC159 in an NF-κB-luciferase reporter assay [82]. In line with previous reports [80-82], expression of wildtype MC159 together with the NF-κB-luciferase reporter plasmid showed a slight increase of NF-κB constitutive activity and effectively blocked TNF-mediated NF-κB activation (Fig. 4 in **III**). However, when $MC159^{AXXA(N)}$ was tested in the same experimental setting, no significant difference in modulating NF-κB activity could be observed, suggesting MC159-SH3-binding not being relevant for this function (Fig.4 in **III**)

To study the inhibition of FAS and TFNR1-induced apoptosis by MC159, we transiently expressed either control plasmid, wildtype MC159 or $MC159^{AXXA(N)}$ in cycloheximide-sensitized HeLa cells, which were 24 hours post-transfection either treated with 10 ng/ml TNF or left untreated. We monitored the Induction of apoptosis by analysing the proteolytic cleavage of caspase-8 and -3 by Western blotting. The reduction of caspase-3 and the appearance of cleaved caspase-8 were detected and quantified. While the quantified signals readily proved the previously reported anti-apoptotic function of wildtype MC159, its SH3 binding capacity did not show any significant effects on this activity, since $MC159^{AXXA(N)}$ inhibited TNF-α induced apoptosis as effectively as wildtype MC159 (Fig. 5 in **III**).

MC159 and other FLIPs have been reported as negative regulators of cellular autophagy [71]. In contrast, SH3BP4 was independently found to possess pro-autophagic functions [91]. Therefore, we investigated the effects of wildtype MC159 and the SH3BP4 binding-deficient $MC159^{AXXA(N)}$ on cellular autophagy. For

this purpose, we used a previously characterized MCF7-derived macroautophagy reporter cell line, which stably expressed enhanced green fluorescent protein-conjugated LC3 (MCF-7/LC3-EGFP) [128] and transduced them with lentiviral vectors encoding for wildtype MC159- or MC159$^{AXXA(N)}$-mCherry fusion proteins. Comparable expression patterns and levels could be observed for stably expressing wildtype MC159 and MC159$^{AXXA(N)}$-mCherry MCF-7/LC3-EGFP cells (Fig. 6A to C in **III**). Importantly, no differences in average number and size of autophagosomes were observed between the two MCF-7/LC3-EGFP cell lines stably expressing MC159 proteins and the parental cell line under standard cell culture conditions (Fig. 6A in **III**).

When amino acids, together with other nutrients, were withdrawn from the parental MCF-7/LC3-EGFP cells for 6 hours, a significant increase in the number of autophagosomes could be detected (Fig. 6B in **III**). In line with the study from Lee and colleagues [71], the MCF-7/LC3-EGFP cells expressing wildtype MC159 showed a less dramatic increase in the number of autophagosomes in response to nutrient starvation (Fig. 6B in **III**). Remarkably, in cells which stably expressed the SH3BP4 binding-deficient MC159$^{AXXA(N)}$, a robust induction of autophagy, similar as seen in the parental cells, could be observed (Fig. 6B in **III**). The Image analysis and quantification of autophagosomes were done using Imaris Software 9.2 from Bitplane, from three independent experiments (Fig. 6E in **III**).

To confirm our findings, we carried out an alternative approach addressing cellular autophagic activity, independent of fluorescent imaging. For this purpose, we analysed intracellular p62 levels by Western blotting, which we obtained from lysates of the same MCF-7/LC3-EGFP parental and stably expressing MC159 cell lines. P62 is a selective autophagy receptor, which delivers its cargo via binding to LC3 to maturing autophagosomes. In the course of autophagic degradation of the delivered cargo, p62 becomes degraded itself and thereby functions as an informative reporter of cellular autophagy [129, 130]. As expected, subjecting the parental cell line to 6 hours nutrient starvation provoked a substantial reduction of detectable p62 protein levels (Fig. 6D in **III**), indicating a robust induction of autophagy in these cells. In agreement with the results obtained from the imaging analysis, cells which stably expressed wildtype MC159 showed, compared to parental cells, a less dramatic fold reduction of p62 levels. This finding supported the previous observation of wildtype MC159 being a negative modulator of autophagy (Fig. 6B & E in **III**). No anti-autophagic activity was observed in SH3BP4-binding deficient MC159$^{AXXA(N)}$ expressing cells, indicated by the profound reduction of p62 protein level (Fig. 6D in **III**).

To prove the direct, functional involvement of SH3BP4 in the observed MC159-SH3BP4 binding-correlated autophagy inhibition, we generated an SH3BP4

knock-out version of the MCF-7/LC3-EGFP cell line by using CRISPR/Cas9 technology. From this knock-out cell line, we generated lentivirally transduced populations, stably expressing either wildtype MC159 or MC159$^{AXXA(N)}$. The absence of SH3BP4, as well as the equal expression of MC159 proteins from the generated cell lines, was confirmed by Western blotting (Fig. 6C in III). The analysis and quantification of autophagosomes were carried out as described for the SH3BP4 expressing MCF-7/LC3-EGFP cells. We observed no changes in the basal number of autophagosomes under standard cell culture conditions when SH3BP4 was absent. However, under starvation conditions, autophagy was slightly lessened, which could be explained by the previously identified role of SH3BP4 as a negative regulator of mTORC1 signalling [91]. Strikingly, unlike in SH3BP4 expressing cells, wildtype MC159 was incapable of counteracting starvation-induced autophagy in cells being depleted of SH3BP4. Inline, the previously observed difference between wildtype MC159 and MC159$^{AXXA(N)}$, functioning as an autophagy inhibitor, was completely absent in SH3BP4 knock-out cells (Fig. 6E in **III**).

Based on the aforementioned data, we concluded that the ability of MCV MC159 protein to counteract starvation-induced autophagy depends on its SH3 domain-mediated interaction with host cell SH3BP4.

5. Discussion

Since the discovery of the SH3 protein domain almost 30 years ago [25], the exploration and understanding of SH3-mediated and -related interactions, as well as their effects on cellular signalling, has greatly advanced our understanding of cellular biology, disease development, and also exploration of the host cell machinery by pathogens. Phage display affinity screens have throughout the years strongly contributed in shedding light into the cellular SH3-interaction network by uncovering novel interactions and as important, revealed how numerous pathogens hijack this network. The discussed results contribute essential and novel aspects of phage display-based identification of SH3 domain-mediated interactions as well as their functional consequences.

A major objective of study **I** was the systematic exploration and mapping of the human "SH3 interactome". Specifically, we aimed for the identification of exceptionally strong and selective SH3 domain-mediated interactions. Following this approach, we intended to understand how frequently SH3 domains play a dominant role in protein interaction partnerships and to discover novel interactions as leads for new studies addressing cell signalling networks.

When screening for novel protein interactions, phage-display technology offers a profound advantage over other screening methods, which is the use of full-length ligand proteins. This feature provides all physiological contacts/ surfaces which are possibly needed for a specific SH3 domain interaction. To ensure the native folding and correct post-translational modification of the potential SH3 ligand proteins, we expressed the selected polypeptides in mammalian HEK293T cells. Using this approach, we were able to express proteins from 324 of the 449 ORFs, which were previously identified as potential SH3 ligands from the human ORFeome library 3.1 by bioinformatical screenings.

Previous studies successfully used the SH3 phage library to identify novel interaction partners for cellular and pathogen-originated ligand proteins and revealed this method as an efficient tool, especially to discover high-affinity SH3 domain-mediated interactions [16, 29, 127, 131, 132]. In this context, a convincing number of identified interactions was subsequently biochemically and functionally verified, supporting the physiological relevance of interactions identified by phage-display screens. One important factor that clearly supports the high success rate in predicting physiologically meaningful SH3-mediated interactions is the feature of the method to uncover high-affinity interactions. An illustrating example is the discovery of the HIV-1 Nef protein interaction with Hck- and Lyn-SH3 domain by phage-display affinity screen [127]. Both

interactions exhibit affinities in the submicromolar range [Nef-Hck-SH3 K_D = 250 nM [121]]. By contrast, the SH3 domains of Lck and Fyn, which have been previously characterized as functional Nef-interaction partners by isothermal titration calorimetry [133], could not be identified by phage-display affinity screening, most likely due to their lower affinities towards Nef (Lyn-SH3 = 10.6 µm and Fyn-SH3 = 15.8 µm [133]. This circumstance exemplifies a technical limitation of the SH3 phage-display in detecting moderate or low-affinity interactions, which nevertheless exhibit biological relevance. Since most SH3-mediated interactions are believed to be of rather transient and thereby low-range affinity nature, it would be certainly of interest to consider measures, which improve the phage-display towards the detection of lower affinity range SH3-mediated interactions. Increasing the SH3-phage library titer or depletion of highly dominant binders from the library could serve as such measures. On the other hand, detecting only high-affinity interactions decreases the false-positive discovery rate significantly, supporting the presented, novel SH3 domain-mediated interactions as likely being biologically relevant.

However, given that out of 324 potential SH3-ligands, only 19 (5.6%) showed defined preferences towards specific SH3 domains, the overall appearance of such strong and selective interactions seems to be rather rare in the human SH3-interaction landscape. It should be noted though, that in total 44 (13.6%) of the 324 ligand proteins showed substantial enrichment of phages over control, suggesting them as functional SH3-ligands. Nevertheless, the remaining 86% of the tested potential SH3-ligands failed to identify as such in this approach. These findings reflect the common understanding that the majority of SH3-mediated interactions are transient and exhibit rather low affinities [134]. Accordingly, novel functional or other non-affinity-based screens should be developed in order to identify such weak but important interactions systematically.

A crucial fact which should not be missed when evaluating SH3 domain-mediated interactions is that proteins and especially proteins which contain modular domains- such as the SH3 domain, mainly exist as components of protein complexes. In this environment, additional interaction surfaces, either between the SH3 interaction partners or provided by other factors of the complex, can significantly increase the overall affinity of these interactions. Since we used isolated and purified ligands in the phage-display screening, SH3-mediated interactions, which depend on such additional contacts, would clearly be missed. The same is true for interactions which rely on the specific localization to certain cellular compartments or active co-localization of the interaction partners. In this context should be further considered that some of the SH3-binding motifs in the screened 324 ligand-proteins are eventually non-accessible and only get exposed by conformational changes of the ligand (e.g. upon its activation by cellular

signalling cascades, alteration in phosphorylation patterns or other post-translational modifications).

Consideration of these aspects sheds new light onto the identified strong and selective SH3 domain-mediated interactions since the two interaction partners alone can establish such highly affine and specific interactions in a cell-free environment. This is further supported by the results we obtained from the semi-quantitative pull-down assay, which suggest for these interactions K_D values in a submicromolar range. Furthermore, they demonstrate that the studied ligand proteins bind their preferred SH3 domain with high selectivity and discriminate even among closely related SH3 domains. Thereby, it is reasonable to assume that highly affine SH3 domain interactions can act as driving forces in defining protein-protein interaction partnerships, also in the context of multiprotein complex assembly. There are indeed examples where the structural analysis of such complexes has proven the previous assumptions, for example, p47hox and p67hox-SH3 [135], PAK and βPIX-SH3 [136], SLP-76 and GADS-SH3 [137] and others. Typical for these interactions are additional areas/ determinants in the ligand-proteins, which increase the overall affinity. These determinants either surround the core PxxP SH3-binding motif, but can also be the result of more complex structural features of the ligand-protein.

The results from the peptide array support the assumption that the high-affinity SH3-mediated interactions, which were identified in our study by phage-display, most likely rely also on such extended contacts. Only a few of the SH3-phage library identified interactions could be recapitulated by the results from the 65 isolated peptides used in the array and vice versa. However, a better overlap of the phage-display and peptide-array results can be seen, when we look at relative binding preferences of the peptides towards specific SH3 domains rather than absolute binding strengths. However, our results suggest that for the identification of SH3-mediated interactions, the use of large-scale peptide-array screens is, due to the peptides' lack of accessory interacting determinants, of limited applicability.

When considering the physiological importance of the newly identified high-affinity interactions from this study, it is essential to note that the strong association of two proteins does not equal a significant biological relevance per se. However, it is quite likely to often be so, since SH3-mediated interactions which were previously identified by our laboratory using this phage-display screening method [127], were identified independently by others [138, 139]. That supports our presented dataset being a valuable and useful resource for further exploration of biological consequences of these newly identified SH3 domain-mediated interactions.

A recent investigation, led by Sachdev Sidhu [12] used a combined approach of random peptide library phage display and deep sequencing. Using this methods, they identified several novel non-canonical SH3-binding motifs. This broad, high-throughput screening approach generated an individual peptide binding-profile for 115 SH3 domains from 89 proteins. It revealed that about half of the examined 115 SH3 domains prefer (at least in this in vitro setting) so far undescribed non-canonical SH3-binding sites. In addition to the well-established class I and II SH3-binding consensus motifs, the study proposes seven new binding motif classes. While the "classes III to VIII" represent potentially novel consensus sequences, "class IX" is a summary of peptide sequences preferred by 35 individual SH3 domains. It would be of interest to extend our biostatistical screening of the human ORFeome library for potential SH3-ligands by these newly identified SH3-binding sequences. Excluding the PxxDY motif (preferred by Eps8(L)1-SH3 domain), which was already included in our initial bioinformatic screen. Of note, this peptide-phage display screen has similar limitations as our peptide-array. The potential screened SH3 domain interaction partners are solely 12 amino acid long peptides, and potentially important additional interaction surfaces from the full-length ligand are missing. Therefore, it is reasonable to assume that despite the comprehensive analysis of the human SH3 interactome by us and others [12], there are most likely several SH3 domain interaction motifs and modes, which have yet to be identified.

Interestingly, the described phenomenon of high-affinity SH3-mediated interactions is frequently described for viral ligand proteins. In the case of HIV-1 Nef binding to Hck-SH3, the exceptional high affinity of the interaction enables the viral protein to dictate the activation state of the non-receptor tyrosine kinase Hck and thereby enhances its kinase activity [140]. Another illustrating example is the SH3-mediated binding of Influenza A virus NS1 protein to Crk adaptor proteins [33], which forces the reorganization of the complex consisting of Crk(L) and P85β towards a trimeric complex in which NS1 is the connecting factor (study II). Furthermore, this interaction can drive the re-localization of Crk(L) from the cytoplasm to the nucleus, which causes changes in the phosphorylation profile of nuclear proteins [141].

Strikingly, and in line with our hypothesis, a very recent study obtained by monitoring the fluorescence intensity change upon complexation, a K_D-value for the NS1 proline-rich motif (PRM)-CrkII-SH3 interaction of 6 nM [142]. This would imply this interaction is about 40 times stronger than the one between Nef and Hck-SH3 and more affine than any other known SH3-mediated interaction. In studies which were previously undertaken in our laboratory, we were able to directly compare HIV-1 Nef and Influenza A NS1 SH3-binding by phage-display (unpublished data). We found Nef-associated phage-enrichment notably

increased compared to the enrichment observed for NS1, which is not in complete agreement with the remarkably low K_D-value reported by [142]. Nevertheless, phage-enrichment of NS1 compared to control was significant and highly selective towards CrkII- and CrkL-SH3, which represented 90% of the identified clones [33]. Thus, and based on the arguments as we presented in study **I** for the newly identified "strong and selective" SH3 domain-mediated interactions, implying a high affinity (in the submicromolar range) for this interaction.

As mentioned above, the highly affine and selective interaction between avian and Spanish flu Influenza A virus NS1 proteins is driving the reorganization of a naturally existing complex between Crk(L) proteins and the PI3K regulatory subunit p85β. This reorganization and the newly formed trimeric complexes, as well as their functional consequences in PI3K signalling, are subject of study **II**. This investigation was initiated after the discovery in our laboratory that avian and Spanish flu Influenza A virus NS1 proteins bind highly selectively to Crk adaptor proteins and as a consequence potentiate PI3K signalling compared to Crk(L)-binding deficient NS1 proteins [33]. To elucidate the mechanism how Crk(L)-binding NS1 proteins are capable of hyper-activating PI3K signalling was of particular interest since all three proteins- NS1, Crk(L) and p85β were found to interact with each other in previous studies. Furthermore, each individual interaction affected PI3K signalling reportedly [33, 56, 57, 143, 144]. Interestingly, the interaction of Crk proteins with the PI3 kinase regulatory subunit p85 was reported to be mediated via the N-terminal SH3 domain of Crk(L) and a proline-rich motif in p85 [57, 126, 143, 145]. Since the same domain was identified by our laboratory to be responsible for the interaction between Crk(L) and SH3-binding competent NS1 proteins [33], we developed the working hypothesis for study **II** that the viral ligand possibly induces a rearrangement of the pre-existing protein Crk(L)-p85 complex.

Using mutational analysis, we identified that the direct interaction of NS1 and p85β, via the critical tyrosine residue in position 89 of NS1, is essential for the hyper-activation of PI3K signalling by Crk(L)-binding competent NS1 proteins. By employing numerous variants of the studied interaction partners, we were able to show that in the case of SH3-binding competent NS1 proteins, the hypnotised reorganization of the pre-existing Crk(L)-p85 complex takes place. In this newly formed "Complex I" NS1 acts as a bridging factor between Crk(L) and p85. Furthermore, we found, that in the presence of SH3-binding deficient NS1 proteins an alternative trimeric complex consisting of Crk(L), p85 and NS1 forms. However, this "Complex II", in which p85 functions as the connecting element, could only be detected when p85 was overexpressed. We observed a clear correlation between the complex architecture and PI3 kinase activity, whereas

complex I promoted an increased PI3 kinase activity. Subsequently, we hypothesized that in the case of complex II, the low abundance of endogenous p85 as well as its moderate affinity towards Crk(L)-SH3 domain, recruited Crk(L) less effectively to the complex and thereby less potently activated PI3 kinase. In favour of this hypothesis, when we overexpressed Crk proteins in cells co-expressing SH3 binding-deficient deficient NS1 proteins, we observed increased pAkt levels, comparable to those in cells co-expressing SH3-binding competent NS1 proteins.

In summary, our findings promote the assumption that the efficacy of Crk protein recruitment to the trimeric complex is determining the extent of PI3 kinase activation. How exactly Crk proteins promote PI3 kinase activation in this context, has still to be elucidated. There are several scenarios imaginable. Since Crk proteins are prototype adaptor molecules, they harbour in addition to their SH3 domain(s) an SH2 domain. Consequently, it is a likely scenario that Crk proteins direct the entire complex by binding to another protein, e.g. via the SH2-domain, to a certain cellular compartment or site where PI3 kinase can efficiently be activated or has improved access to its substrates. The findings that the interaction between Crk proteins and NS1 is highly affine [142] and that NS1 has recently been found to cause a relocation of Crk proteins from the cytosol to the nucleus [141], support this hypothesis. Furthermore, there are reported Crk interaction partners who could function as such docking sites for Crk proteins and are known to locally promote PI3 kinase activation, namely the focal adhesion kinase (FAK) [145], p130cas [146] and c-Cbl [147].

Since the basic mechanism of PI3 kinase activation by NS1 proteins is the release of inhibitory contacts between p85 regulatory- and p110 catalytic PI3 kinase subunits [144], another reasonable role of Crk(L) in PI3 kinase activation could be an additional contribution in destabilizing inhibitory contacts. However, the highly efficient binding of NS1 to Crk proteins and the subsequently increased recruitment of the adaptor proteins to the PI3K complex, represents most likely a key element in avian and Spanish flu NS1 PI3 kinase hyper-activation.

Notably, the same study which recently reported the remarkable K_D value for the NS1 PRM and CrkII-SH3 interaction, identified additional sites, adjacent to the PxLP core SH3-binding motif in NS1, which significantly contributed to the strong interaction. Especially the combination of short-range electrostatic interactions between the NS1 PRM lysine 217 and the specificity area of Crk-SH3 domain (D147, E149 and D150) as well as the formation of hydrogen bonds between two NS1 PRM amide backbones (Q218, K219) and Crk-SH3 glutamic acid 166, promoted the interaction. Furthermore, the overall positive charge of the NS1 PRM establishes long-distance electrostatic interactions, which support the

observed rapid association of NS1 with Crk proteins [142]. This extended interaction surface enables NS1 to hijack Crk proteins from their regular cellular interaction partners and to manipulate further signalling pathways which are regulated by Crk(L), as we could show in study **II**.

The functional importance of NS1-Crk(L)-SH3 binding for IAV life-cycle seems to be, however, strongly strain- and subsequently host-dependent [148]. While the NS1 SH3-binding motif appears to be important for avian IAV replication in their natural hosts [125], it does in most cases not provide any benefit when it is introduced to NS1 proteins which naturally lack the binding site [149]. As mentioned before, IAV strains which naturally lack Crk(L)-binding ability, are mainly adapted to humans. While the primary site of replication in the human host are epithelial cells of the respiratory tract, in birds, IAV replicates in intestinal epithelial cells. Consequently, the cellular environment in which human and avian IAV replicate is significantly different, and the need for specific host-cell factors is likely to differ as well. In this context, intermediate hosts susceptible to IAV strains of both origins, human and avian, are of particular interest for host-cell factor requirements of the virus. Hale and colleagues have tested, whether the introduction of a functional Crk(L)-binding site (E217 to K mutation) in NS1 enhances viral replication of the 2009 pandemic H1N1 IAV in human, but also swine cells. The results showed no increase in viral replication in either cell type [147]. When considering the recent identification of the adjacent binding sites in the NS1 PRM, that contribute significantly to the affinity between NS1 and Crk(L)-SH3, one could certainly argue that a single amino acid change does not create the necessary high affinity. Yet, a single amino acid change in IAV Udorn NS1 (T215 to proline) created a Crk(L)-SH3 binding site identical to the one, which exhibited the K_D value of 6 nM towards CrkII-SH3 [142]. Another explanation could be that swine tracheal epithelial cells are physiologically too close to those from humans to show an effect when introducing the Crk(L)-SH3 binding site into NS1. In recent years, chickens and their role as a potential intermediate host gained more attention. Like in pigs, chicken tracheal epithelial cells exhibit sialic acid linked to galactose by α-2,6 linkages (SAα2,6Gal), preferably bound by human IAV and SAα2,3Gal, bound by avian IAV. Notably, SAα2,6 Gal dominates in the chicken trachea, rendering them susceptible infection with IAV from both origins [150]. Yet, the physiology of these cells is avian-like and thereby typical avian IAV host factors like Crk(L) are likely to be indispensable, eventually also for strains which do not naturally contain an SH3 binding motif. Consequently, it could be of interest to test human IAV strains in chickens, which express SH3 binding-competent NS1 proteins.

A recent investigation led by Benjamin G. Hale analysed the evolution of NS1-p85β binding of human IAV strains, which circulated globally during the last

century [151]. The experimental results indicate that enhanced binding of NS1 to p85β is advantageous for IAV in competition assays. Their thorough bioinformatic analysis of over 24000 NS1 sequences also revealed that NS1 variants harbouring residues, which promoted increased p85β binding, occurred and were maintained in the human population over decades in the 20th century. However, also NS1 variants with reduced p85β binding capacity appeared and established in human IAV strains since the end of the 20th century [151]. The authors propose that a decrease in viral fitness, caused by reduced p85β binding, could have been compensated by changes in other parts of the virus. Our findings from study **II** shows that IAV has also developed other tools, namely the SH3 domain-mediated binding to Crk proteins, to increase PI3kinase signalling. Nevertheless, it is crucial to note that the study from Turkington and colleagues addressed only the correlation between NS1-p85β binding with viral fitness/ survival and not the effect of PI3kinase activation in this context. Another remarkable finding from the same laboratory is that NS1 proteins from two bat IAV isolates, are entirely incapable of binding p85β and activating PI3 kinase [151]. This discovery underlines the extensive adaptation of IAV to their hosts, and thereby the diversity of the different IAV strains in hijacking host cell signalling factors and pathways.

The phenomenon of viruses being highly adapted to their host's cell environment is a key strategy of viral survival. Since IAV encodes for a relatively small number of proteins, it is not too surprising that the NS1 protein interacts with and manipulates host cellular factors in such a versatile way. In contrast, the viral protein which is subject of study **III**, MCV MC159, is part of a large pox-viral machinery consisting of 182 proteins. Also, this protein has, similar to IAV NS1, shown to be a key immuno-evasion factor of the virus. Though, it seems at first glance somewhat surprising that a single protein is bearing such essential functions when the virus expresses such a battery of viral proteins. Considering the life-cycles of IAV and MCV, it becomes more obvious why the latter requires a complex network of viral factors tightly controlling the host cell. MCV, alike many large DNA-viruses, establishes long-lasting infections which can persist up to 4 years [63]. Therefore MCV has developed to be almost invisible to the host cellular immune system, and this requires tools precisely containing anti-viral actions of the host [64]. MC159 is together with MC160, one of the two viral FLIPs expressed by MCV. Viral FLIPs and their cellular counterparts are pro-caspase-like proteins, which function as key regulators of cellular apoptosis [74, 75, 152], and inhibitors of cellular autophagy [71].

Apart from being a key factor in viral immune evasion, MC159 and IAV NS1 have another feature in common; they both comprise bona fide SH3-binding sites. In fact, this feature is unique to MC159 among the FLIP family and made it of

particular interest to our laboratory. Starting our work on MC159, we identified two SH3-binding consensus motifs, a class I consensus motif in the very N-terminus of the protein and a class II motif in the C-terminal proximity. However, we could only identify the N-terminal class I motif as a functional SH3-binding site using the phage-display affinity screen, which revealed its exclusive binding to the SH3 domain of cellular SH3BP4. When looking at the second class II motif in MC159, it is puzzling that no preferred SH3 domain could be identified. The motif contains, apart from the typical key residues promoting SH3-binding (PxϕPx+), additional flanking residues possibly contributing to higher affinities in that context [142]. As already discussed in the context of study **I**, it is important to keep in mind that the identification of SH3-interaction partners by phage-display also has limitations. However, selective binding to the SH3 domain of SH3BP4 could be observed for the N-terminal class I motif, making it the key determinant of MC159 acting as SH3-ligand. The location of the motif in the viral protein is, however, unusual. In many viral SH3-ligand proteins the binding sites are located in unstructured, highly flexible regions, as it is the case in IAV NS1 [33]. In contrast, in MC159 the motif is part of the DED1 helices H0 and H1. Notably, in the crystal structure of MC159 the key prolines at positions eight and 11 point in opposite directions [72], suggesting a conformational change in MC159 as either required for prior or happening upon binding of the SH3BP4 SH3 domain. This could also explain the high preference of the motif towards SH3BP4 SH3 domain, despite its classical consensus sequence, indicating the binding site as rather promiscuous.

At the time, we identified SH3BP4 as an interaction partner of MC159, only limited information about this cellular protein was available. However, the few articles which were published suggested the involvement of SH3BP4 in critical cellular processes. SH3BP4 was found to regulate transferrin receptor internalization by promoting cargo-specific clathrin-mediated endocytosis [93]. It was also found to be critical for FGF-receptor trafficking after internalization and subsequent functional events [94]. However, of particular relevance to study **III** was the finding that SH3BP4 is involved in the regulation of cellular autophagy. By directly binding to Rag-GTPases, SH3BP4 prevents their recruitment to mTORC1 and thereby inhibits amino acid-induced mTORC1 activity. This consequently induces autophagy and further inhibits cellular growth [91].

Since MC159, among other viral FLIPs, had independently been shown to inhibit cellular autophagy [71], the effect of MC159 binding to SH3BP4 in the context of autophagy induction was of particular interest for us. Moreover, the pro-autophagic activity of SH3BP4 was reported to be dependent on the same SH3 domain, which we found to be responsible for MC159-SH3BP4 interaction [91].

In fact, our results showed that loss of the functional N-terminal SH3-binding site, rendered MC159 being incapable of counteracting starvation-induced autophagy. Furthermore, we found that in cells which were depleted of SH3BP4, the existence of a functional SH3-binding site became irrelevant, and neither wildtype nor the AxxA-variant of MC159 could inhibit starvation-induced autophagy.

When envisioning the mechanism of how MC159 counteracts autophagy by binding to SH3BP4 SH3 domain, the most obvious scenario would be that MC159 prevents SH3BP4 from binding to Rag-GTPases and thereby supports mTORC1 activity. However, it was shown before that MC159 also counteracts the mTORC1-inhibitor rapamycin, which acts independently of the Rag-GTPase-mTORC1 association status [71]. Moreover, the same study identified that viral and cellular FLIPs inhibit autophagy downstream of mTORC1 by binding the E2-like enzyme Atg3 and thereby prevent the processing of LC3I to LC3II. However, it is important to note that a causal relation between vFLIP-Atg3 binding and autophagy inhibition was only shown for KSHV FLIP K13. In this context, the authors found the K13 H2 helix of DED1 and H4 helix of DED2 being indispensable for the reported effect. Despite MC159 reportedly showed an Atg3 interaction profile similar to KSHV vFLIP, the DED1 of both proteins have considerable structural differences [72], which makes a direct transfer of functional effects of KSHV vFLIP on MC159 questionable. Also, considering the mechanistic diversity in vFLIP-mediated inhibition of death receptor-induced apoptotic signalling [64], it is reasonable to assume that also in the context of autophagy inhibition, different strategies are employed. Since MCV encodes for a second vFLIP, MC160, it would be informative to test, whether this protein also shows Atg3-binding activity and inhibitory effects on autophagy. Nevertheless, the investigation of a possible effect or contribution of MC159-SH3BP4 interaction on the binding of Atg3 to MC159 should receive further attention.

As discussed, binding of MC159 to SH3BP4 causes possibly conformational changes in MC159, which could also have consequences on the ability of MC159 binding to other proteins, e.g. Atg3. This suggested conformational change also allows envisioning so far unconsidered mechanisms of how MC159-SH3BP4 interaction contributes to autophagy inhibition, independent of Rag-GTPases or Atg3. For example, MC159 harbours a potential so-called LC3-interacting region (LIR) motif “SRFVEL” starting at position 139 in its second DED [153]. Under isolated conditions, the DEDs of MC159 adopt a rigid dumbbell-shaped conformation, in which the valine 142 of the LIR motif is completely buried [72]. If a conformational change in MC159 would release both DEDs from each other, the motif could become functional. LIR motifs have a fundamental role in selective autophagy by mediating the binding of autophagy receptors (e.g. p62)

to phagophore-anchored Atg8 proteins (e.g. LC3) and thereby allow the delivery of the receptor-cargo towards autophagosomal degradation [154]. As a result, binding of MC159 to Atg8 proteins could block the binding of selective autophagy receptors and their cargo-delivery.

However, of the previously reported MC159 functions we examined in study **III**, autophagy remained the only one, on which alteration of MC159-SH3BP4 interaction had a measurable effect. In this context, wildtype and the SH3BP4 binding-deficient AxxA variant showed comparable activities in moderately elevating basal levels of NFκB-activity, but also similar capacities in profoundly inhibiting TNF-induced NFκB-activity. Further, the loss of SH3BP4-binding ability neither showed any effect on TNF-stimulated apoptosis, since both- wildtype MCF159 and its AxxA variant potently inhibited proteolytic caspase-8 and -3 cleavage. Conclusively, at this point, the selective binding of MC159 to SH3BP4 evolved seemingly to support inhibition of autophagy.

In addition to the already mentioned functions of SH3BP4 [93-95], more recent findings support the importance of this protein in various additional cell signalling pathways. Just lately, Antas et al. reported that SH3BP4 negatively regulates Wnt-signaling by modulating the nuclear localization of β-catenin. Thereby SH3BP4 functions as tumour-suppressor in intestinal stem cells [97]. Another study from Luo et al. further links SH3BP4 to amino acid regulated mTORC1-signaling. The authors state that upon amino acid starvation of cow mammary gland epithelial cells (CMECs), sestrin2 (SESN2) is expressed and binds to SH3BP4, what leads to negative regulation of mTORC1 as described by Kim et al. [95] and subsequently to reduced synthesis of casein [155].

Interestingly, SH3BP4 has also been reported to be part of a multiprotein complex, called “LIFT” which is thought to regulate endosome positioning and -intracellular movement by remodelling of actin [96]. Myosin VI is a key component of this complex and is also known to bind selective autophagy receptors, and by its action, as “actin-based motor” it helps to deliver the bound receptor-cargo complexes to the growing phagophore [106]. Whether the LIFT-complex, and thereby SH3BP4 and a putatively bound MC159, could also directly be involved in this process would surely be of interest for further investigations. Apart from this, SH3BP4 increasingly gains importance as a biomarker in different cancer types, such as laryngeal and metastatic prostate cancer [156, 157].

These novel findings add not only to the understanding of SH3BP4-biology but also highlight the functional diversity of this protein. Our finding of MCV MC159 specifically targeting SH3BP4 to support its inhibitory effect on cellular autophagy adds yet another piece to the picture, which outlines the complex role

of SH3BP4 in the cell. Further investigations addressing the question of whether MC159-SH3BP4 interaction also affects other functions of SH3BP4 will be of great interest. Preliminary work of our laboratory on the alteration of TfR1-internalization by binding of MC159 to SH3BP4 showed an effect in transferrin-uptake assays. However, a direct correlation of this effect with SH3BP4-binding capacity of MC159 could not be shown, suggesting an indirect function of MC159 underlying this observation (C.Schmotz, unpublished data).

In summary, the results of study **III** clearly support the hypothesis that recruitment of SH3BP4 by MC159 N-terminal SH3-binding motif actively contributes to its ability to suppress autophagy. MCV establishes persisting infections which can last for years. For this purpose, the virus has developed complex strategies which make it highly adapted to its host cells, human keratinocytes, and almost invisible to the host's immune system. The suppression of host cellular autophagy is knowingly one of these strategies and is typically employed by large DNA-viruses establishing long-lasting infections [112].

6. Conclusions

“As the elucidation of the structure of the DNA double helix was to the molecular biology, so the discovery of modular binding domains was important in reorienting and focusing ideas in the signalling and opening the door to new experimental approaches” [5]. This statement by Professor Bruce Mayer, who is one of the pioneers in the field of modular protein domains, emphasises the fundamental role of modular domains in cellular signalling. SH3 domains represent the archetype and most abundant group of protein modular binding domains. The human genome encodes for more than 300 different SH3 domains, expressed in 220 proteins, demonstrating their absolute necessity for the cell.

In general, SH3 domain-mediated interactions are considered to exhibit rather moderate affinities, mostly because of their participation in transient protein complex assemblies. By reporting the discovery of numerous novel, high-affinity SH3 domain-mediated interactions, we illustrate another, commonly rarely appreciated feature of this vital protein networking module. Moreover, the presented findings clearly support phage-affinity screen technology as a bona fide method for the identification of such strong and selective SH3 domain interactions. By direct comparison of results we obtained from phage-screenings, peptide-arrays and semi-quantitative pull-down assays, we can conclude that a critical aspect in high affinity and high selectivity of SH3 domain interactions is the existence of additional contacts between the ligand-protein and the preferred SH3 domain. Despite our finding that this type of high-affinity SH3-mediated interaction is rather the exception than the rule in the human SH3 interactome, the novel interactions we identified, are a valuable resource for future investigations addressing functional consequences or structural features.

The relevance and impact of highly affine and selective SH3 domain-mediated interactions becomes further evident in the context of host-cell signalling exploration by pathogens, especially viruses. Our findings shed light into how IAV NS1 hijacks and manipulates host PI3K signalling by establishing a strong interaction with cellular Crk adaptor proteins and thereby forcing the reorganization of existing host cellular protein interactions. The findings prove how powerful NS1 interferes with elements of host cellular signalling, by adoption, incorporation and probably further improvement of an SH3 binding motif. Nevertheless, we still lack an answer to the question, how IAV benefits from hyper-activation of PI3K signalling by Crk(L)-binding competent NS1 proteins. The interplay of the host cell- and IAV strain origin clearly plays a significant role in this context. However, future work on this topic is required to answer this question fully.

Of no less significance is the finding of MCV MC159 protein recruiting SH3BP4 via an SH3-mediated interaction and the discovery that this viral FLIP depends on binding to host cell SH3BP4-SH3 domain to inhibit the cellular autophagy machinery efficiently. Our findings add thereby essential pieces to the big puzzle of understanding, how this virus escapes the human immune defence almost completely. However, there is undoubtedly still a lot to learn about this multifunctional key factor of MCV immune evasion and its actions. The elucidation of the mechanism behind autophagy inhibition, supported by MC159 binding to SH3BP4, to begin with.

In conclusion, the presented results provide an outline of the presence and frequency of high affinity and selective SH3-mediated interactions in the human SH3 interactome. In that context, we identified strong and selective SH3 partners for 19 potential SH3-ligand proteins and thereby provided a useful platform for future investigations. The results we present on the reorganization of the Crk(L)-PI3K-signalling complex by SH3-binding competent IAV NS1 proteins, underline the consequences such high affinity and selective SH3-mediated interactions can have on host cell protein complex formation and signalling when hijacked by pathogens. The data we present on MCV MC159 protein binding to host cell SH3BP4, introduces a new member into the group of viral SH3 ligands. Moreover, by identifying this interaction as an essential feature of MC159 for its efficient suppression of cellular autophagy, we introduce another example how such highly selective and strong interactions of viral SH3 domain ligands with their target SH3 domain can manipulate host cell signalling.

7. Acknowledgments

This study was carried out at the Department of Virology, Medicum, University of Helsinki. I want to thank the head of the department, Professor Kalle Saksela, for providing excellent research facilities as well as a welcoming and inspiring working environment.

I want to genuinely thank and express my gratitude to my supervisor, Professor Kalle Saksela, for his support and patience during this journey. Your scientific knowledge and experience are outstanding. Thank you for sharing it with me.

I warmly thank Docent Varpu Marjomäki and Professor Marku Pesu for reviewing this thesis and for all the valuable comments I received, they improved the work substantially. I sincerely would like to thank my thesis committee members Associate Professor Denis Kainov and Docent Sampsa Matikainen for the accompaniment and evaluation of this thesis throughout the years.

Further, I sincerely want to thank Adjunct Professor Petri Susi for accepting the role as the opponent in my thesis defence and Professor Olli Vapalahti for taking on the position as the faculty representative.

I gratefully thank all collaborators and co-authors, without whom this thesis would not have been possible. Leena Ylösmäki, Erkko Ylösmäki, Arunas Kazlauskas, Riku Fagerlund, Tapio Kesti, Silja Vilen, Iivari Kleino, Jussi Hepojoki, Hasan Uğurlu, Subhash Shrestha, Tomonori Kaneko and Shawn Li.

This thesis has been partly financed by the Integrative Life Science Doctoral Program (former Graduate Program in Biotechnology and Molecular Biology), University of Helsinki. The Biomedicum Helsinki Foundation and Juhani Aho Foundation for Medical Research are thanked for their financial support.

I sincerely want to thank all current and former members of the Saksela lab. Kalle, Virpi, Leena, Erkko, Annika, Hannamari, Tapio, Silja, Arunas, IIvari, Riku, Tina, Matjaz, Kristina, Anna, Zhao, Hasan, Yoke, Subhash, Johanna, Marcal, Sergio, Jubayer, Inka and Giedrius. Thank you for all your support and for creating such a great working atmosphere. Especially I want to thank Leena, Annika, Virpi and Erkko, not only for teaching me all kind of methods and tricks in the lab but also genuinely for your friendship, help and advice in all aspects of life. Tina and Matjaz, thank you for being such dear friends to me. I also like to thank Riku very much, without your contribution and expertise, the last article wouldn't have been possible.

Hannamari and Pentti, thank you so much for all the conversations we had over the past years, you always encouraged me to keep on going.

A big thank you goes also to all other people from the Department of Virology for lifting up the spirit every day. I am so grateful that among you I found dear friends, to whom I could always turn to and count on, Suvi, Satu K., Erika and Satu H., thank you so much.

The Biomedicum Imaging Unit, especially Mikko Liljestöm, I want to thank for the great support with the image analysis and for putting up with my frequent consultations.

I also warmly thank Professor Sanna Lehtonen and her research group, among whom I found my new “scientific home”.

I like to thank all the people who keep Haartman Institute running every day. Especially Pete, Markku (and before Jorma) and the entire Unicafe staff, who make coming to work every day better. I genuinely want to thank the first two people I met when I arrived at Haartman Institute ten years ago. Our former secretary Hannele and vahtimestari Heikki, who are sadly both not with us anymore. Thank you for making me feel so welcome and taking on the scary pile of bureaucracy waiting for a student moving to a foreign country.

I cannot thank enough my friends here in Finland and back in Germany. My girls, Jenine, Gigi, Vicky, Elba and Christine, from here and Claudi, Jessi and both Anjas from Germany. I am so lucky to have you all in my life. Hanna, Sanna, Katja and Hanni, how lucky can one be and have such wonderful friends right next door.

Words cannot explain how thankful I am to my family, raising me with love and all their support. My mother, grandma, sister and little brother, you always have my back, without you I wouldn't be. Frank and grandpa, we miss you every day. My deepest gratitude goes as well to my family in Spain. Thank you for your warm welcoming and all your help.

With all my heart I want to thank my better half Juan. We have come all this way together, and you never stopped believing in me. Without your support, nothing of this would have been possible.

Last but not least, I want to thank my beloved children Carlotta and Patrik. You are the brightest side of life. The most important lessons I learned from you.

8. References

1. Cesareni, G., et al., *Modular Protein Domains*. 2005, Weinheim, Germany: Wiley-VHC.
2. White, S.H. and R.E. Jacobs, *Statistical distribution of hydrophobic residues along the length of protein chains. Implications for protein folding and evolution.* Biophys J, 1990. **57**(4): p. 911-21.
3. George, R.A. and J. Heringa, *SnapDRAGON: a method to delineate protein structural domains from sequence data.* J Mol Biol, 2002. **316**(3): p. 839-51.
4. Sidhu, S.S., W.J. Fairbrother, and K. Deshayes, *Exploring protein-protein interactions with phage display.* Chembiochem, 2003. **4**(1): p. 14-25.
5. Mayer, B.J., *The discovery of modular binding domains: building blocks of cell signalling.* Nat Rev Mol Cell Biol, 2015. **16**(11): p. 691-8.
6. Salazar, M.A., et al., *Tuba, a novel protein containing bin/amphiphysin/Rvs and Dbl homology domains, links dynamin to regulation of the actin cytoskeleton.* J Biol Chem, 2003. **278**(49): p. 49031-43.
7. Kaneko, T., L. Li, and S.S. Li, *The SH3 domain--a family of versatile peptide- and protein-recognition module.* Front Biosci, 2008. **13**: p. 4938-52.
8. Branden, C. and J. Tooze, *Introduction to Protein Structure*. 2nd ed. 2012: Graland Science. 424.
9. Saksela, K. and P. Permi, *SH3 domain ligand binding: What's the consensus and where's the specificity?* FEBS Lett, 2012. **586**(17): p. 2609-14.
10. Lim, W.A., *Reading between the lines: SH3 recognition of an intact protein.* Structure, 1996. **4**(6): p. 657-9.
11. Feng, S., et al., *Molecular basis for the binding of SH3 ligands with non-peptide elements identified by combinatorial synthesis.* Chem Biol, 1996. **3**(8): p. 661-70.
12. Teyra, J., et al., *Comprehensive Analysis of the Human SH3 Domain Family Reveals a Wide Variety of Non-canonical Specificities.* Structure, 2017. **25**(10): p. 1598-1610.e3.
13. Kurochkina, N. and U. Guha, *SH3 domains: modules of protein-protein interactions.* Biophys Rev, 2013. **5**(1): p. 29-39.
14. Karavasilis, V., et al., *CHAPTER 17 - Cancer drug resistance*, in *Cancer Drug Design and Discovery*, S. Neidle, Editor. 2008, Academic Press: New York. p. 405-423.

15. Smith, K.M., R. Yacobi, and R.A. Van Etten, *Autoinhibition of Bcr-Abl through its SH3 domain.* Mol Cell, 2003. **12**(1): p. 27-37.
16. Kleino, I., et al., *Alternative splicing of ADAM15 regulates its interactions with cellular SH3 proteins.* J Cell Biochem, 2009. **108**(4): p. 877-85.
17. Gouw, M., et al., *The eukaryotic linear motif resource – 2018 update.* Nucleic Acids Research, 2017. **46**(D1): p. D428-D434.
18. Bliska, J., *How pathogens exploit interactions mediated by SH3 domains.* Chem Biol, 1996. **3**(1): p. 7-11.
19. Ren, S., et al., *Short Linear Motifs recognized by SH2, SH3 and Ser/Thr Kinase domains are conserved in disordered protein regions.* BMC Genomics, 2008. **9 Suppl 2**: p. S26.
20. Via, A., et al., *How pathogens use linear motifs to perturb host cell networks.* Trends Biochem Sci, 2015. **40**(1): p. 36-48.
21. Vogt, P.K., *Retroviral oncogenes: a historical primer.* Nat Rev Cancer, 2012. **12**(9): p. 639-48.
22. Feller, S.M., et al., *SH2 and SH3 domains as molecular adhesives: the interactions of Crk and Abl.* Trends Biochem Sci, 1994. **19**(11): p. 453-8.
23. Smart, J.E., et al., *Characterization of sites for tyrosine phosphorylation in the transforming protein of Rous sarcoma virus (pp60v-src) and its normal cellular homologue (pp60c-src).* Proc Natl Acad Sci U S A, 1981. **78**(10): p. 6013-7.
24. Mayer, B.J., M. Hamaguchi, and H. Hanafusa, *A novel viral oncogene with structural similarity to phospholipase C.* Nature, 1988. **332**(6161): p. 272-5.
25. Reichman, C.T., et al., *The product of the cellular crk gene consists primarily of SH2 and SH3 regions.* Cell Growth Differ, 1992. **3**(7): p. 451-60.
26. Fujinami, R.S., et al., *Molecular mimicry in virus infection: crossreaction of measles virus phosphoprotein or of herpes simplex virus protein with human intermediate filaments.* Proc Natl Acad Sci U S A, 1983. **80**(8): p. 2346-50.
27. Saksela, K., G. Cheng, and D. Baltimore, *Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4.* Embo j, 1995. **14**(3): p. 484-91.
28. Nanda, S.K., D. Herion, and T.J. Liang, *The SH3 binding motif of HCV [corrected] NS5A protein interacts with Bin1 and is important for apoptosis and infectivity.* Gastroenterology, 2006. **130**(3): p. 794-809.
29. Neuvonen, M., et al., *SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication.* PLoS Pathog, 2011. **7**(11): p. e1002383.

30. Rom, S., et al., *HIV-1 Tat binds to SH3 domains: cellular and viral outcome of Tat/Grb2 interaction.* Biochim Biophys Acta, 2011. **1813**(10): p. 1836-44.
31. Lim, M.J., et al., *Suppression of c-Src activity stimulates muscle differentiation via p38 MAPK activation.* Arch Biochem Biophys, 2007. **465**(1): p. 197-208.
32. Aitio, O., et al., *Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21743-8.
33. Heikkinen, L.S., et al., *Avian and 1918 Spanish influenza a virus NS1 proteins bind to Crk/CrkL Src homology 3 domains to activate host cell signaling.* J Biol Chem, 2008. **283**(9): p. 5719-27.
34. Bouvier, N.M. and P. Palese, *The biology of influenza viruses.* Vaccine, 2008. **26 Suppl 4**: p. D49-53.
35. Taubenberger, J.K. and D.M. Morens, *The pathology of influenza virus infections.* Annu Rev Pathol, 2008. **3**: p. 499-522.
36. WHO. *Influenza (Seasonal)*. 2019; Available from: https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal).
37. Barberis, I., et al., *History and evolution of influenza control through vaccination: from the first monovalent vaccine to universal vaccines.* J Prev Med Hyg, 2016. **57**(3): p. E115-e120.
38. CDC. *History of 1918 Flu Pandemic*. 2018, March 21; Available from: https://www.cdc.gov/flu/pandemic-resources/1918-commemoration/1918-pandemic-history.htm.
39. Racaniello, V. *Influenza virus RNA genome*. 2009, May 1; Available from: http://www.virology.ws/2009/05/01/influenza-virus-rna-genome/.
40. Te Velthuis, A.J. and E. Fodor, *Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis.* Nat Rev Microbiol, 2016. **14**(8): p. 479-93.
41. Dou, D., et al., *Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement.* Front Immunol, 2018. **9**: p. 1581.
42. CDC, *How the Flu Virus Can Change: "Drift" and "Shift".* 2019, October 15.
43. Hale, B.G., et al., *The multifunctional NS1 protein of influenza A viruses.* J Gen Virol, 2008. **89**(Pt 10): p. 2359-76.
44. Mossman, K., *Viruses and Interferon: Current Research*. 2011: Caister Academic Press.
45. Iwasaki, A. and P.S. Pillai, *Innate immunity to influenza virus infection.* Nat Rev Immunol, 2014. **14**(5): p. 315-28.

46. Randall, R.E. and S. Goodbourn, *Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures.* J Gen Virol, 2008. **89**(Pt 1): p. 1-47.
47. Killip, M.J., et al., *Single-cell studies of IFN-beta promoter activation by wild-type and NS1-defective influenza A viruses.* J Gen Virol, 2017. **98**(3): p. 357-363.
48. Li, S., et al., *Binding of the influenza A virus NS1 protein to PKR mediates the inhibition of its activation by either PACT or double-stranded RNA.* Virology, 2006. **349**(1): p. 13-21.
49. Min, J.Y., et al., *A site on the influenza A virus NS1 protein mediates both inhibition of PKR activation and temporal regulation of viral RNA synthesis.* Virology, 2007. **363**(1): p. 236-43.
50. Ludwig, S., et al., *The influenza A virus NS1 protein inhibits activation of Jun N-terminal kinase and AP-1 transcription factors.* J Virol, 2002. **76**(21): p. 11166-71.
51. Javier, R.T. and A.P. Rice, *Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses.* J Virol, 2011. **85**(22): p. 11544-56.
52. Golebiewski, L., et al., *The avian influenza virus NS1 ESEV PDZ binding motif associates with Dlg1 and Scribble to disrupt cellular tight junctions.* J Virol, 2011. **85**(20): p. 10639-48.
53. Hale, B.G., et al., *Binding of influenza A virus NS1 protein to the inter-SH2 domain of p85 suggests a novel mechanism for phosphoinositide 3-kinase activation.* J Biol Chem, 2008. **283**(3): p. 1372-80.
54. Li, Y., et al., *Mechanism of influenza A virus NS1 protein interaction with the p85beta, but not the p85alpha, subunit of phosphatidylinositol 3-kinase (PI3K) and up-regulation of PI3K activity.* J Biol Chem, 2008. **283**(34): p. 23397-409.
55. Birge, R.B., et al., *Crk and CrkL adaptor proteins: networks for physiological and pathological signaling.* Cell Commun Signal, 2009. **7**: p. 13.
56. Sattler, M., et al., *Steel factor induces tyrosine phosphorylation of CRKL and binding of CRKL to a complex containing c-kit, phosphatidylinositol 3-kinase, and p120(CBL).* J Biol Chem, 1997. **272**(15): p. 10248-53.
57. Gelkop, S., Y. Babichev, and N. Isakov, *T cell activation induces direct binding of the Crk adapter protein to the regulatory subunit of phosphatidylinositol 3-kinase (p85) via a complex mechanism involving the Cbl protein.* J Biol Chem, 2001. **276**(39): p. 36174-82.
58. Mauldin, E.A. and J. Peters-Kennedy, *Chapter 6 - Integumentary System*, in *Jubb, Kennedy & Palmer's Pathology of Domestic Animals: Volume 1 (Sixth Edition)*, M.G. Maxie, Editor. 2016, W.B. Saunders. p. 509-736.e1.

59. Fox, R., et al., *Molluscum contagiosum in two donkeys.* Vet Rec, 2012. **170**(25): p. 649.
60. Chen, X., A.V. Anstey, and J.J. Bugert, *Molluscum contagiosum virus infection.* Lancet Infect Dis, 2013. **13**(10): p. 877-88.
61. Connell, C.O., et al., *Congenital molluscum contagiosum: report of four cases and review of the literature.* Pediatr Dermatol, 2008. **25**(5): p. 553-6.
62. Patrizi, A., et al., *Congenital molluscum contagiosum.* Paediatrics & Child Health, 2017. **22**(5): p. 241-242.
63. Badri, T. and G.R. Gandhi, *Molluscum Contagiosum*, in *StatPearls*. 2019, StatPearls Publishing

StatPearls Publishing LLC.: Treasure Island (FL).
64. Shisler, J.L., *Immune evasion strategies of molluscum contagiosum virus.* Adv Virus Res, 2015. **92**: p. 201-52.
65. Perez-Diaz, C.E., et al., *Giant Molluscum Contagiosum in an HIV positive patient.* Int J Infect Dis, 2015. **38**: p. 153-5.
66. Senkevich, T.G., et al., *Genome sequence of a human tumorigenic poxvirus: prediction of specific host response-evasion genes.* Science, 1996. **273**(5276): p. 813-6.
67. Senkevich, T.G., et al., *The Genome of Molluscum Contagiosum Virus: Analysis and Comparison with Other Poxviruses.* Virology, 1997. **233**(1): p. 19-42.
68. Hendrickson, R.C., et al., *Orthopoxvirus genome evolution: the role of gene loss.* Viruses, 2010. **2**(9): p. 1933-67.
69. Yu, J.W. and Y. Shi, *FLIP and the death effector domain family.* Oncogene, 2008. **27**: p. 6216.
70. Shisler, J.L., *Viral and cellular FLICE-inhibitory proteins: a comparison of their roles in regulating intrinsic immune responses.* J Virol, 2014. **88**(12): p. 6539-41.
71. Lee, J.S., et al., *FLIP-mediated autophagy regulation in cell death control.* Nat Cell Biol, 2009. **11**(11): p. 1355-62.
72. Yang, J.K., et al., *Crystal structure of MC159 reveals molecular mechanism of DISC assembly and FLIP inhibition.* Mol Cell, 2005. **20**(6): p. 939-49.
73. Thurau, M., et al., *The TRAF3-binding site of human molluscipox virus FLIP molecule MC159 is critical for its capacity to inhibit Fas-induced apoptosis.* Cell Death Differ, 2006. **13**(9): p. 1577-85.
74. Bertin, J., et al., *Death effector domain-containing herpesvirus and poxvirus proteins inhibit both Fas- and TNFR1-induced apoptosis.* Proc Natl Acad Sci U S A, 1997. **94**(4): p. 1172-6.

75. Hu, S., et al., *A novel family of viral death effector domain-containing molecules that inhibit both CD-95- and tumor necrosis factor receptor-1-induced apoptosis.* J Biol Chem, 1997. **272**(15): p. 9621-4.
76. Garvey, T., et al., *The death effector domains (DEDs) of the molluscum contagiosum virus MC159 v-FLIP protein are not functionally interchangeable with each other or with the DEDs of caspase-8.* Virology, 2002. **300**(2): p. 217-25.
77. Garvey, T.L., et al., *Binding of FADD and caspase-8 to molluscum contagiosum virus MC159 v-FLIP is not sufficient for its antiapoptotic function.* J Virol, 2002. **76**(2): p. 697-706.
78. Li, F.Y., et al., *Crystal structure of a viral FLIP: insights into FLIP-mediated inhibition of death receptor signaling.* J Biol Chem, 2006. **281**(5): p. 2960-8.
79. Fu, T.M., et al., *Cryo-EM Structure of Caspase-8 Tandem DED Filament Reveals Assembly and Regulation Mechanisms of the Death-Inducing Signaling Complex.* Mol Cell, 2016. **64**(2): p. 236-250.
80. Randall, C.M., J.A. Jokela, and J.L. Shisler, *The MC159 protein from the molluscum contagiosum poxvirus inhibits NF-kappaB activation by interacting with the IkappaB kinase complex.* J Immunol, 2012. **188**(5): p. 2371-9.
81. Gil, J., et al., *MC159L protein from the poxvirus molluscum contagiosum virus inhibits NF-kappaB activation and apoptosis induced by PKR.* J Gen Virol, 2001. **82**(Pt 12): p. 3027-34.
82. Murao, L.E. and J.L. Shisler, *The MCV MC159 protein inhibits late, but not early, events of TNF-alpha-induced NF-kappaB activation.* Virology, 2005. **340**(2): p. 255-64.
83. Liu, T., et al., *NF-κB signaling in inflammation.* Signal Transduction And Targeted Therapy, 2017. **2**: p. 17023.
84. Solt, L.A. and M.J. May, *The IkappaB kinase complex: master regulator of NF-kappaB signaling.* Immunol Res, 2008. **42**(1-3): p. 3-18.
85. Biswas, S. and J.L. Shisler, *Molluscum Contagiosum Virus MC159 Abrogates cIAP1-NEMO Interactions and Inhibits NEMO Polyubiquitination.* J Virol, 2017. **91**(15).
86. Randall, C.M., et al., *Inhibition of interferon gene activation by death-effector domain-containing proteins from the molluscum contagiosum virus.* Proc Natl Acad Sci U S A, 2014. **111**(2): p. E265-72.
87. Dunlevy, J.R., et al., *Cloning, chromosomal localization, and characterization of cDNA from a novel gene, SH3BP4, expressed by human corneal fibroblasts.* Genomics, 1999. **62**(3): p. 519-24.
88. Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

89. Khanobdee, K., J.B. Kolberg, and J.R. Dunlevy, *Nuclear and plasma membrane localization of SH3BP4 in retinal pigment epithelial cells.* Mol Vis, 2004. **10**: p. 933-42.
90. Mayer, B.J., *Endocytosis: EH domains lend a hand.* Curr Biol, 1999. **9**(2): p. R70-3.
91. Kim, Y.M. and D.H. Kim, *dRAGging amino acid-mTORC1 signaling by SH3BP4.* Mol Cells, 2013. **35**(1): p. 1-6.
92. EBML-EBI, *ZU5 domain.* 2017.
93. Tosoni, D., et al., *TTP specifically regulates the internalization of the transferrin receptor.* Cell, 2005. **123**(5): p. 875-88.
94. Francavilla, C., et al., *Functional proteomics defines the molecular switch underlying FGF receptor trafficking and cellular outputs.* Mol Cell, 2013. **51**(6): p. 707-22.
95. Kim, Y.M., et al., *SH3BP4 is a negative regulator of amino acid-Rag GTPase-mTORC1 signaling.* Mol Cell, 2012. **46**(6): p. 833-46.
96. O'Loughlin, T., T.A. Masters, and F. Buss, *The MYO6 interactome reveals adaptor complexes coordinating early endosome and cytoskeletal dynamics.* EMBO Rep, 2018. **19**(4).
97. Antas, P., et al., *SH3BP4 Regulates Intestinal Stem Cells and Tumorigenesis by Modulating beta-Catenin Nuclear Localization.* Cell Rep, 2019. **26**(9): p. 2266-2273.e4.
98. Yu, L., Y. Chen, and S.A. Tooze, *Autophagy pathway: Cellular and molecular mechanisms.* Autophagy, 2018. **14**(2): p. 207-215.
99. Parzych, K.R. and D.J. Klionsky, *An overview of autophagy: morphology, mechanism, and regulation.* Antioxid Redox Signal, 2014. **20**(3): p. 460-73.
100. English, L., et al., *Autophagy enhances the presentation of endogenous viral antigens on MHC class I molecules during HSV-1 infection.* Nat Immunol, 2009. **10**(5): p. 480-7.
101. Svenning, S. and T. Johansen, *Selective autophagy.* Essays Biochem, 2013. **55**: p. 79-92.
102. Hou, W., et al., *Strange attractors: DAMPs and autophagy link tumor cell death and immunity.* Cell Death Dis, 2013. **4**: p. e966.
103. Mancias, J.D. and A.C. Kimmelman, *Mechanisms of Selective Autophagy in Normal Physiology and Cancer.* J Mol Biol, 2016. **428**(9 Pt A): p. 1659-80.
104. Ying, H. and B.Y. Yue, *Optineurin: The autophagy connection.* Exp Eye Res, 2016. **144**: p. 73-80.
105. Zheng, L., et al., *Recent Advances in Understanding Amino Acid Sensing Mechanisms that Regulate mTORC1.* Int J Mol Sci, 2016. **17**(10).

106. Kast, D.J. and R. Dominguez, *The Cytoskeleton-Autophagy Connection.* Curr Biol, 2017. **27**(8): p. R318-r326.
107. Axe, E.L., et al., *Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum.* J Cell Biol, 2008. **182**(4): p. 685-701.
108. Monastyrska, I., et al., *Multiple roles of the cytoskeleton in autophagy.* Biol Rev Camb Philos Soc, 2009. **84**(3): p. 431-48.
109. Levine, B., *Eating oneself and uninvited guests: autophagy-related pathways in cellular defense.* Cell, 2005. **120**(2): p. 159-62.
110. Deretic, V., T. Saitoh, and S. Akira, *Autophagy in infection, inflammation and immunity.* Nat Rev Immunol, 2013. **13**(10): p. 722-37.
111. Romao, S., M. Gannage, and C. Munz, *Checking the garbage bin for problems in the house, or how autophagy assists in antigen presentation to the immune system.* Semin Cancer Biol, 2013. **23**(5): p. 391-6.
112. Choi, Y., J.W. Bowman, and J.U. Jung, *Autophagy during viral infection - a double-edged sword.* Nat Rev Microbiol, 2018. **16**(6): p. 341-354.
113. Jackson, W.T., et al., *Subversion of cellular autophagosomal machinery by RNA viruses.* PLoS Biol, 2005. **3**(5): p. e156.
114. Prentice, E., et al., *Coronavirus replication complex formation utilizes components of cellular autophagy.* J Biol Chem, 2004. **279**(11): p. 10136-41.
115. Saksela, K. and D. Baltimore, *Negative regulation of immunoglobulin kappa light-chain gene transcription by a short sequence homologous to the murine B1 repetitive element.* Mol Cell Biol, 1993. **13**(6): p. 3698-705.
116. Melen, K., et al., *Nuclear and nucleolar targeting of influenza A virus NS1 protein: striking differences between different virus subtypes.* J Virol, 2007. **81**(11): p. 5995-6006.
117. Frank, R., *The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications.* J Immunol Methods, 2002. **267**(1): p. 13-26.
118. Lamesch, P., et al., *hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes.* Genomics, 2007. **89**(3): p. 307-15.
119. Faravelli, A. and N. Dimasi, *Expression, refolding and crystallizations of the Grb2-like (GADS) C-terminal SH3 domain complexed with a SLP-76 motif peptide.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2006. **62**(Pt 1): p. 52-5.
120. Takemoto, Y., et al., *Distinct binding patterns of HS1 to the Src SH2 and SH3 domains reflect possible mechanisms of recruitment and activation of downstream molecules.* Int Immunol, 1996. **8**(11): p. 1699-705.

121. Lee, C.H., et al., *A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein.* Embo j, 1995. **14**(20): p. 5006-15.
122. Lee, C.H., et al., *Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain.* Cell, 1996. **85**(6): p. 931-42.
123. Hale, B.G., et al., *Influenza A virus NS1 protein binds p85beta and activates phosphatidylinositol-3-kinase signaling.* Proc Natl Acad Sci U S A, 2006. **103**(38): p. 14194-9.
124. Brehme, M., et al., *Charting the molecular network of the drug target Bcr-Abl.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7414-9.
125. Hrincius, E.R., et al., *Avian influenza viruses inhibit the major cellular signalling integrator c-Abl.* Cell Microbiol, 2014. **16**(12): p. 1854-74.
126. Sattler, M., et al., *The proto-oncogene product p120CBL and the adaptor proteins CRKL and c-CRK link c-ABL, p190BCR/ABL and p210BCR/ABL to the phosphatidylinositol-3' kinase pathway.* Oncogene, 1996. **12**(4): p. 839-46.
127. Karkkainen, S., et al., *Identification of preferred protein interactions by phage-display of the human Src homology-3 proteome.* EMBO Rep, 2006. **7**(2): p. 186-91.
128. Hoyer-Hansen, M., et al., *Control of macroautophagy by calcium, calmodulin-dependent kinase kinase-beta, and Bcl-2.* Mol Cell, 2007. **25**(2): p. 193-205.
129. Klionsky, D.J., et al., *Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition).* Autophagy, 2016. **12**(1): p. 1-222.
130. Komatsu, M., et al., *Homeostatic levels of p62 control cytoplasmic inclusion body formation in autophagy-deficient mice.* Cell, 2007. **131**(6): p. 1149-63.
131. Asbach, B., et al., *Comprehensive analysis of interactions between the Src-associated protein in mitosis of 68 kDa and the human Src-homology 3 proteome.* PLoS One, 2012. **7**(6): p. e38540.
132. Kesti, T., et al., *Reciprocal regulation of SH3 and SH2 domain binding via tyrosine phosphorylation of a common site in CD3epsilon.* J Immunol, 2007. **179**(2): p. 878-85.
133. Arold, S., et al., *RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef.* Biochemistry, 1998. **37**(42): p. 14683-91.
134. Ladbury, J.E. and S. Arold, *Searching for specificity in SH domains.* Chem Biol, 2000. **7**(1): p. R3-8.
135. Kami, K., et al., *Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p.* Embo j, 2002. **21**(16): p. 4268-76.

136. Hoelz, A., et al., *Crystal structure of the SH3 domain of betaPIX in complex with a high affinity peptide from PAK2.* J Mol Biol, 2006. **358**(2): p. 509-22.
137. Harkiolaki, M., et al., *Structural basis for SH3 domain-mediated high-affinity binding between Mona/Gads and SLP-76.* Embo j, 2003. **22**(11): p. 2571-82.
138. Yuan, Z.Q., et al., *ArgBP2gamma interacts with Akt and p21-activated kinase-1 and promotes cell survival.* J Biol Chem, 2005. **280**(22): p. 21483-90.
139. Yuan, Z.Q., et al., *ArgBP2gamma interacts with Akt and p21-activated kinase-1 and promotes cell survival.* J Biol Chem, 2016. **291**(43): p. 22845.
140. Moarefi, I., et al., *Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement.* Nature, 1997. **385**(6617): p. 650-3.
141. Ylosmaki, L., et al., *Nuclear Translocation of Crk Adaptor Proteins by the Influenza A Virus NS1 Protein.* Viruses, 2016. **8**(4): p. 101.
142. Shen, Q., et al., *The Molecular Mechanisms Underlying the Hijack of Host Proteins by the 1918 Spanish Influenza Virus.* ACS Chem Biol, 2017. **12**(5): p. 1199-1203.
143. Akagi, T., et al., *v-Crk activates the phosphoinositide 3-kinase/AKT pathway in transformation.* Proc Natl Acad Sci U S A, 2000. **97**(13): p. 7290-5.
144. Hale, B.G., et al., *Structural insights into phosphoinositide 3-kinase activation by the influenza A virus NS1 protein.* Proc Natl Acad Sci U S A, 2010. **107**(5): p. 1954-9.
145. Akagi, T., et al., *v-Crk activates the phosphoinositide 3-kinase/AKT pathway by utilizing focal adhesion kinase and H-Ras.* Mol Cell Biol, 2002. **22**(20): p. 7015-23.
146. Zhang, J., et al., *CRKL Mediates p110beta-Dependent PI3K Signaling in PTEN-Deficient Cancer Cells.* Cell Rep, 2017. **20**(3): p. 549-557.
147. Husson, H., et al., *CSF-1 stimulation induces the formation of a multiprotein complex including CSF-1 receptor, c-Cbl, PI 3-kinase, Crk-II and Grb2.* Oncogene, 1997. **14**(19): p. 2331-8.
148. Ayllon, J., et al., *Contribution of NS1 effector domain dimerization to influenza A virus replication and virulence.* J Virol, 2012. **86**(23): p. 13095-8.
149. Hale, B.G., et al., *Mutations in the NS1 C-terminal tail do not enhance replication or virulence of the 2009 pandemic H1N1 influenza A virus.* J Gen Virol, 2010. **91**(Pt 7): p. 1737-42.
150. Nidom, C.A., et al., *Influenza A (H5N1) viruses from pigs, Indonesia.* Emerg Infect Dis, 2010. **16**(10): p. 1515-23.

151. Turkington, H.L., et al., *Unexpected Functional Divergence of Bat Influenza Virus NS1 Proteins.* J Virol, 2018. **92**(5).
152. Irmler, M., et al., *Inhibition of death receptor signals by cellular FLIP.* Nature, 1997. **388**(6638): p. 190-5.
153. Jacomin, A.C., et al., *iLIR database: A web resource for LIR motif-containing proteins in eukaryotes.* Autophagy, 2016. **12**(10): p. 1945-1953.
154. Birgisdottir, A.B., T. Lamark, and T. Johansen, *The LIR motif - crucial for selective autophagy.* J Cell Sci, 2013. **126**(Pt 15): p. 3237-47.
155. Luo, C., et al., *Sestrin2 Negatively Regulates Casein Synthesis through the SH3BP4-mTORC1 Pathway in Response to AA Depletion or Supplementation in Cow Mammary Epithelial Cells.* J Agric Food Chem, 2019. **67**(17): p. 4849-4859.
156. Wilkins, O.M., et al., *MicroRNA-Related Genetic Variants Associated with Survival of Head and Neck Squamous Cell Carcinoma.* Cancer Epidemiol Biomarkers Prev, 2019. **28**(1): p. 127-136.
157. Zhang, X., et al., *Investigation of the molecular mechanisms underlying metastasis in prostate cancer by gene expression profiling.* Exp Ther Med, 2016. **12**(2): p. 925-932.